Science Library

OCT 21 1948

Vol. 38, No. 5

October, 1948

Journal of Experimental Psychology

EDITED BY

FRANCIS W. IRWIN

UNIVERSITY OF PENNSYLVANIA

CONSULTING EDITORS

CLARENCE H. GRAHAM, *Columbia University*
HARRY HELSON, *Bryn Mawr College*
DONALD B. LINDSLEY, *Northwestern University*
DONALD G. MARQUIS, *University of Michigan*
GEORGE A. MILLER, *Harvard University*
L. L. THURSTONE, *University of Chicago*
G. R. WENDT, *University of Rochester*
JOSEPH ZUBIN, *N. Y. Psychiatric Institute, Columbia University*

CONTENTS

Postremity, Recency, and Frequency as Bases for Prediction in the Maze Situation: VIRGINIA W. VOEKS 495

On the Problem of 'Reinforcement' in Conditioning the Autokinetic Phenomenon: E. A. HAGGARD AND RACHEL RABIN 511

Further Study of the Retention of Verbal and Motor Skills: FRANCES VAN DUSEN AND H. SCHLOSBERG 526

Rate of Recall as a Measure of Learning: I. The Effects of Retroactive Inhibition: L. POSTMAN, J. P. EGAN, AND JEAN DAVIS 535

The Sequential Order of Concept Attainment: BERNICE M. WENZEL AND CHRISTINE FLURRY 547

Double Alternation Behavior in Young Children: W. S. HUNTER AND SUSAN C. BARTLETT 558

The Effect on Group Performance of an Indifferent and Neglectful Attitude Shown by One Group Member: D. ROSENTHAL AND C. N. COFER 568

The Relative Sweetness of Sugars: Sucrose and Dextrose: P. E. LICHTENSTEIN 578

The Rate of Blinking during Prolonged Visual Search: A. CARPENTER 587

Apparent Movement in Relation to Homonymous and Heteronymous Stimulation of the Cerebral Hemispheres: J. A. GENGERELLI 592

Infant Speech: Speech Sound Development of Sibling and Only Infants: O. C. IRWIN 600

Studies in Space Orientation. III. Perception of the Upright in the Absence of a Visual Field: H. A. WITKIN AND S. E. ASCH 603

Studies of Tracking Behavior. I. Rate and Time Characteristics of Simple Corrective Movements. L. V. SEARLE AND F. V. TAYLOR 615

PUBLISHED BI-MONTHLY BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

PRINCE AND LEMON STS., LANCASTER, PA.

AND 1515 MASSACHUSETTS AVE., N. W., WASHINGTON 5, D. C.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879

Acceptance for mailing at the special rate of postage provided for in the Act of February 28, 1925, embodied in paragraph 4, Section 538, P. L. & R., authorized September 11, 1947.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published bi-monthly by the American Psychological Association, Inc. The annual subscription rate is $7.00 (foreign, $7.25), single copies, $1.25. Subscriptions, orders, and business communications should be addressed to the American Psychological Association, Inc., 1515 Massachusetts Ave., N.W., Washington 5, D. C.

This JOURNAL publishes articles which are intended to contribute toward the development of psychology as an experimental science. Experimental work with normal human subjects is favored over work with abnormal or animal subjects. Articles of a length exceeding 25 printed pages cannot be accepted.

With rare exceptions, articles are published in the order of their receipt. Authors are supplied with 50 free offprints, without covers, in the case of articles of seven or more printed pages, and with five free copies of the issue in the case of shorter contributions. One-half of the cost of tables, cuts, formulas, and author's alterations in galley proof is charged to the author.

Immediate publication can be given to articles whose authors assume the full cost of publication.

Address articles submitted for publication to the editor, Francis W. Irwin, Department of Psychology, University of Pennsylvania, Philadelphia 4, Pa.

This JOURNAL adheres to the conventions concerning bibliographies, citations, preparation of tables and cuts, etc., as described by J. E. ANDERSON and W. L. VALENTINE, The preparation of articles for publication in the journals of the American Psychological Association, *Psychol. Bull.*, 1944, **41**, 345–376. Special attention is called to the style of numbers and abbreviations (p. 353). The term 'experimenter,' with its various forms, singular, plural, and possessive singular and plural, is abbreviated *E*, *E*s, *E*'s, and *E*s'; analogous abbreviations are used for the terms 'subject' (*S*) and 'observer' (*O*). All manuscripts should be typed double-spaced throughout, *including tables, footnotes, and bibliography*. Tables, figures, and footnotes should be placed on separate pages at the end of the manuscript.

Journal of
Experimental Psychology

VOL. 38, NO. 5 OCTOBER, 1948

POSTREMITY, RECENCY, AND FREQUENCY AS BASES FOR PREDICTION IN THE MAZE SITUATION [1]

BY VIRGINIA W. VOEKS

University of Washington

INTRODUCTION

A central problem in psychology is the determination of bases for accurate prediction of behavior, with the necessity thus arising of ascertaining what factors determine the response an organism will make to any given situation. There have been postulated a variety of such factors, the present investigation being conducted to determine the significance of one of these: namely, the factor of recency in the Guthrian sense of the term (cf. **3**, pp. 417 and 420, **4**, p. 457, **5**, **6**, p. 527, **7**, p. 26, **8**, p. 111). The following is a brief statement of this principle of recency: In a recurring situation, such as repeated trials in a puzzle-box or maze, the most probable response on any trial is the one last made to the stimuli present on that trial. If some transitory cause alters the response, this new response is the one predicted for the next trial on which that stimulus pattern occurs. This advantage of the last-made response holds even when some other response has had indefinitely greater frequency.

Although there have been many studies of 'recency,' most of these were concerned with principles other than the present one. For example, a great number centered around the effect of differences in the length of time interval between presentation of a situation and re-presentation of that situation, or the effect of differences in position of a pair of nonsense syllables in a list. None of this work has any direct bearing on the present problem of recency. However, because of the duplication in terms, conclusions regarding 'recency'

[1] This paper is taken from a dissertation presented to the faculty of the University of Washington in partial fulfillment of the requirements for the Master of Science degree.

reached by those experimenters have been widely quoted by others discussing the newer 'recency' principle, even though the studies and conclusions were not about the same principle and the experimental data on those principles show nothing about the present principle of recency. Partly because of a tendency to ignore differences in meaning and implication, and partly because of other related factors, considerable confusion has resulted from using the same term to cover a variety of meanings.[2] To help preclude further confusion, it would be well to introduce a new term for at least one of these principles of recency. Accordingly in the present paper, the term 'postremity' (from the Latin *postremo*, superlative of 'late') will be used when referring to the Guthrian principle, and the term 'recency' when referring to the more traditional principles.

As previously mentioned, most of the experimental work on recency has no significance regarding postremity; but some experiments have been done to investigate this latter principle. In 1937, Guthrie and Horton (9) observed the behavior of cats in a puzzle-box, release from which was effected by tilting a pole. Having a semispherical base, this pole could be tilted easily by any one of a great variety of responses. A photograph was taken automatically of each escape movement, and it was found that "although no two cats used the same method, the successive escapes of any individual cat showed an unexpected number of final movements substantially identical." A detailed account of this experiment is presented in the 1946 book by these co-workers. A further study by Horton and Guthrie (11) of cats in a puzzle-box of considerably different construction revealed again a marked tendency toward stereotyped escape movements. As the principle of postremity would lead one to predict, the cats tended strongly to do in this situation what they had last done in that situation.

For the maze situation, Peterson is the only experimenter who has made an effort to test the validity of the principle of postremity. (He called it 'recency.') Two experiments were performed. In the first (17), 17 rats were divided into three groups, each group learning one of three differently patterned mazes. For every rat, a record was kept of the specific blinds entered on any given trial; but the results for the first three trials only were published. It is unfortunate that only the first three trials were analyzed, for during the beginning trials especially the situation is changing rapidly and rather markedly as the internal pattern of stimuli is often greatly modified by, for example, the dispersion of initial excitement, curiosity, or tension. Thus on the first few trials the possible validity of the principle of postremity is somewhat obscured.

The conclusion drawn by Peterson from this study was that while "the effects of recency and frequency are obvious in any detailed study" (p. 360), they are not of any great importance. Nonetheless, it is to be noted that for each group of rats, and for 12 of the 17 individuals, the predictions made

[2] For various meanings of 'recency,' see the usages of Brown, Hunter, Jersild, McGeoch, Warden and Cummings, Watson.

on the basis of postremity were more often correct than incorrect. Such results would occur by chance about five times in 100. The lack of greater statistical significance may be attributable to the small number of rats being used and the first trials only being analyzed.

In a second experiment (18), 16 human *Ss* learned a 10-blind, mental maze, the results of all trials being published in this work. Peterson concluded that postremity is not a factor in learning, and stated further that it is "negative in its effects, tending to fix the erroneous response rather than the correct one" (p. 290). These conclusions, although widely quoted, lack statistical justification. The data show that of 4,518 predictions based upon postremity, 68.4 percent were correct. This degree of accuracy differs from 50.0 percent by 26.1 times the standard error of 50.0 percent (SE = .74 percent). From this rough analysis it would seem that postremity of a response is a sound basis for prediction of behavior.[3]

Furthermore, upon analysis of the data, postremity is not found to be negative in its effect, but on the contrary, is found to fix apparently the right responses more than the erroneous ones. Of those occasions when the postreme response had been the 'right' response, 73.0 percent of the time it was made on the next trial. Of those occasions when the postreme response had been the 'wrong' response, only 63.9 percent of the time was it made on the next trial. Therefore, it appears that if postremity tends more toward preserving one set of responses rather than another, it tends more toward preserving the correct responses rather than the incorrect. (Actually, by the present hypothesis, postremity itself is neutral, having no differential effect on the correct and incorrect responses.) Thus it is found that contrary to Peterson's conclusion, his data indicate that postremity may be a significantly accurate basis for prediction and is not negative in its effect.

Yoshioka (23) used 21 rats in a three-legged maze, each of these legs making an angle of 120° with the others. The rat was started at the end of one, and was fed at the end of the next leg to which he went. He then was started from the end of that leg, and fed at the end of that to which he next went. After this he was removed from the maze. For each rat, a record was kept of the responses on each set of trials. Yoshioka concluded that "preferential pattern may be one of the dominant factors in maze learning" (p. 197), but that frequency and recency are negligible. It should be noted that Yoshioka was not using 'recency' in the sense of postremity. In fact, he meant by 'pattern' what partially is meant herein by postremity. If, for example, a rat went from the end of one of the legs, turned right at the central choice point, and proceeded to the end of that second leg, Yoshioka would predict on the basis of 'preferred pattern' that when the rat went from the end of that leg and arrived next at the central choice point, he would again turn right. This is precisely what would be predicted on the basis of postremity. The rat, having run down a straight path and having then found two other paths extending at a certain angle toward the right and toward the left, last responded by turning toward the right. There-

[3] This statistical analysis is not strictly correct. For example, it in no way takes into account individual differences, nor does it equally weight the results from each subject. As the data for individuals were not published, it was impossible to overcome these and other difficulties.

fore, when next the rat runs down a straight path and is again at the central choice point with a path extending at the same angle toward the right and one toward the left, it would be predicted from postremity (if the other stimuli are similar) that he will turn right, for his postreme response to such a stimulus pattern was turning toward the right. The empirical findings strongly indicate the validity of this basis for prediction, the number of times the responses accorded with predictions based upon 'preferential pattern' or postremity being significant at the one percent level of confidence.

Hudson (12) also performed an experiment in which one-trial learning clearly was demonstrated. Using 14 rats, he once shocked the animals as they ate from a metallic food-holder fastened to a star design. After one month, he again put the animals in that situation and found that 13 of the 14 made conditioned responses of withdrawal to that stimulus pattern. According to our calculations based on the binomial expansion, this result would occur by chance less than one time in a thousand.

Other experiments could be analyzed to determine the significance (or non-significance) of the principle of postremity; but the data from most of these were lumped, all errors for a given trial being grouped or data from all subjects being grouped, and thus the possibility of such an analysis of these data is precluded.

Statement of the Problem

The present experiment was designed to aid in answering these specific questions: (1) Can postremity be used as a basis for accurate predictions of behavior? (2) If the postreme response is one which had been made *less* frequently in the past than some other response, is postremity an accurate basis for prediction? (3) What is the rellative accuracy of postremity and frequency as bases for prediction?

Apparatus and Procedure

In order to have a fairly adequate indication of the generality of any results which might be obtained, three parallel experiments were conducted, in each of which different subjects with different apparatus and procedures were used. The experiments were conducted in a quiet and relatively isolated, but unfortunately not soundproof, room.

In the first experiment, a high-relief, 18 choice-point, multiple-T, finger maze was used in which the end of each blind was marked by a small brad, and the end of each correct alley was marked by a slight hump in the wire immediately preceding the choice point. This was patterned after a maze designed by Spence (19), with the modifications introduced by Granneberg (2). With a practice maze of this type before him, the procedure was explained to each *S* and he was told that he would be required to reach a criterion of three successive perfect trials on the regular maze. Detailed instructions were given for *S*'s counting of alleys (*S* paused after his finger dropped down from the hump and announced the number of the choice point *before* that choice was made), for the position of the hand and fingers, and for the complete procedure. The instructions were accompanied by a demonstration by *E*, and were followed by *S*'s closing his eyes and going through the practice maze once. He was then blindfolded, the maze to be learned was placed before him, and some of the directions were repeated for emphasis and for the purpose of giving him an opportunity to become accustomed to the blindfold. Care was taken in all cases to have the subject feel at ease before the experiment was begun.

For the second experiment, a punch-board stylus maze was used. This maze had 30 pairs of holes in a figure 'S' pattern identical to that of the maze introduced by Muenzinger and Vine (16).

A chime sounded if the *S* chose the correct hole of a pair; a buzzer, if he chose the incorrect one. The pattern of correct and incorrect holes was reversed at the end of each *S*'s trials, so that the holes which had been correct for the last *S* were incorrect for the next *S* and those which had been incorrect were correct. These *S*s were instructed also to stop and say aloud the number of each pair before making their choice, if an error was made to notice it carefully but not to correct it at that time, to continue working on the maze until the criterion of three perfect successive trials was reached, and to "go at a moderate rate; do *not* rush through the maze, but do not make long stops in any part of it—don't, for example, when you get to number 4, try to rehearse your moves that far as that will cause you to do more poorly." [4]

In the third experiment, one change only was made: The *S*s were given a threaded stylus which necessitated the screwing of the stylus into the holes and out again after every choice, whereas in the second experiment the stylus was merely pushed in.

The 57 *S*s were volunteers from classes in Elementary Psychology at the University of Washington. For the first experiment, 25 *S*s were used (Group 1); for the second, 16 (Group 2); and for the third, 16 (Group 3).

Method of Recording Data and Making Predictions

For each individual, a record of responses was kept on graph paper, the response made on each trial at each choice point being recorded in the proper square for that trial and choice point. For each response after trial 1, two predictions were made: One prediction was made on the basis of postremity—the response recorded for the preceding trial and that choice point being the response predicted for the present trial. The other prediction was made on the basis of frequency—the response recorded most often on all preceding trials for that choice point being the one predicted for the present trial. If the two opposing responses had been made an equal number of times at that point, either of the two would be predicted by frequency. For such cases, one-half the time the 'correct' response was recorded for the prediction and one-half the time the 'incorrect,' the predictions being alternated for those points.

Results and Interpretation

A. Analysis of all predictions made by postremity.

Considering the data as a whole, predictions based upon postremity were found to agree extremely well with observed behavior. For Group 1, the mean of the percentage correct for each *S* is 88.7 percent; for Group 2, 82.3 percent; for Group 3, 83.1 percent

TABLE I

Total Predictions by Postremity for Groups

	Number of Subjects	Mean % Correct Predictions	Sigma of		CR
			Dist.	Mean	
Group 1	25	88.74	2.97	0.606	63.9
Group 2	16	82.30	4.09	1.06	30.4
Group 3	16	83.05	4.45	1.15	28.7

* In Tables I and II, *N* is taken as the number of *S*s; the means are obtained by adding the percentage correct for each *S* in the group and dividing by the number of *S*s; sigmas of the mean equal sigma of the distribution of scores divided by $\sqrt{N-1}$; the critical ratio equals the mean minus 50 percent divided by the sigma of the mean.

[4] Much implicit practice was avoided by permitting no pauses between trials, nor any pauses of more than three sec. during the trials.

(Table I). The critical ratios of these means to their standard errors are 63.9, 30.4, and 28.7 respectively.

Also, examination of the record for each of the 57 individuals separately revealed that *for every S the predictions based upon postremity were highly accurate.*[5] For each of 49 *S*s, more than $\frac{4}{5}$ of the predictions made by postremity were correct. For only two of 57 individuals were less than $\frac{3}{4}$ of the predictions correct, and for only one of these did postremity predict correctly as little as 73.3 percent of the time. The CR of 73.3 percent minus 50.0 percent to the SE of 50.0 percent is 6.2; so it is seen that even for that individual for whom postremity was least accurate, the accuracy was high enough that, for this number of trials, it would not be equalled by chance one time in a million.[6] The results for all other individuals also were significant considerably beyond the 0.1 percent level.

B. Analysis of those Predictions by postremity which were opposed by frequency.

For this analysis, not all the predictions made by postremity were considered, but only those for which the response postremity predicted was not the one that had been more frequently made. It was found that the mean percentages of these predictions which were made correctly by postremity were about 71 percent for the various groups, with the CR's being about 8 (see Table II).

TABLE II

PREDICTIONS BY POSTREMITY WHEN AT VARIANCE WITH FREQUENCY FOR GROUPS

	Number of Subjects	Mean % Correct Predictions	Sigma of		CR
			Dist.	Mean	
Group 1	25	68.91	11.9	2.43	7.8
Group 2	16	68.10	5.58	1.44	12.6
Group 3	16	76.56	7.89	2.04	8.2

An analysis was made also of the results for each individual. In the cases of 24 of the 25 members of Group 1, even though postremity was predicting the response which had been made less frequently in the past, its predictions were correct the major part of the time. Such a result would occur by chance about three times in a million.

[5] Tabulations of these data for each individual and other individual data mentioned in sections A, B, and C (number of responses predicted correctly, number predicted incorrectly, standard errors, critical rations, *P*-values, etc.) were omitted to conserve space. They may be obtained from the writer.

[6] The formula used for this and each of the other standard errors for individuals in sections A and B is: $\sqrt{50\,(50)/N}$, in which N is the number of predictions made for that individual.

Similarly in Group 2, postremity predicted correctly more often than incorrectly in the case of each of the 16 *Ss*. If the null hypothesis were correct, all 16 cases would fall in the same direction about three times in 100,000. The same results were obtained for Group 3, 100 percent of those 16 cases also being in the same direction.

Moreover, even for these instances in which postremity predicted the less frequent response, it predicted correctly more than $\frac{2}{3}$ of the time in the cases of each of 41 *Ss* and more than $\frac{3}{4}$ of the time for 23 *Ss*. The predictions by postremity were significant beyond the five percent level of confidence for 46 *Ss*, beyond the one percent level for 37 *Ss*, and beyond the 0.1 percent level for 29 *Ss*.

It is apparent that for the great majority of individuals and for all groups in this experiment, postremity is a statistically reliable basis for prediction of responses, *even when frequency is favoring an opposed response.*

C. Comparison of the predictions by postremity and those by frequency when at variance.

The final analysis was a study of the relative accuracy of postremity and frequency as bases for predictions when the two are at variance. As can be seen from Table II, for Group 1, 68.91 is the mean percentage of postremity's predictions which were correct when an opposed response had been more frequent; whereas only 31.09 percent of frequency's predictions were correct on the average when it predicted a response other than the postreme response. Similar results were found for Groups 2 and 3, predictions based upon frequency having a mean accuracy of only 31.90 and 23.44 percent when postremity was opposed.

If, instead of grouping the data, each *S*'s record is examined separately, the findings are comparable. When at variance with postremity, for only seven *S*'s out of 57 were predictions by frequency correct as much as 40 percent of the time; and for 41 *Ss* they were correct *less* than $\frac{1}{3}$ of the time. A statistical analysis showed that for each of 47 individuals, the greater accuracy of postremity's predictions over those of frequency would not occur five times in 100 by chance; for 37, not once in 1,000.[7]

Also, for 56 of the 57 individuals used in all groups, when opposed by postremity, frequency predicted with less accuracy than postremity, and with less than 50 percent accuracy. If the principle of postremity held with no consistently greater accuracy than that of frequency, these results would not occur once in 10 million groups of 57 individuals.

[7] The formula used for the SE of the differences followed Treloar (20, p. 227).

Thus we find that in this experiment frequency is an accurate basis of prediction only as long as predictions based upon it agree with predictions based upon postremity. When the more frequent response is not the postreme response, then frequency is a highly inaccurate basis of prediction, less accurate than chance.

Further Discussion

From the foregoing analysis, it is apparent that to a marked degree in at least some maze situations, predictions based upon the principle of postremity correspond to observed behavior. There remain, however, a number of questions in the literature concerning its interpretation and implications.

A. Variations in response.

One of the main questions is this: If the principle of postremity is valid, why does an animal not continue forever to make the same response at any given choice point?

Part of the answer is that for a large proportion of its training, the animal does do precisely that: If he once happens to make the correct response, he then makes the correct response repeatedly; if he once makes an error at a choice point, he continues to make that same error repeatedly. In the present experiment, more than $\frac{4}{5}$ of the responses were of this nature, were repetitions of the response last made.

But the *S* does not always make the same response when at the same choice point. Why is this? One of the fundamental factors causing a change in response at a given point is that there are marked changes from trial to trial in the pattern of stimuli present—even at the 'same' choice point.

In the first place, there are changes from one trial to the next in the internal stimulus pattern. Especially is this so during the first few trials, for during them the *S*, rat or human, may be quite emotional and excited, and the resultant pattern of internal stimuli may change fairly rapidly. But later in training also, the internal stimuli may change considerably from trial to trial: for example, human *S*s may develop feelings of elation, depression, impatience, or anxiety as the trials progress, thus changing the visceral stimuli. Likewise, there are changes in energy level and muscle tonus, and changes as the result of fatigue.

Moreover, as the result of imperfect control, there are further shifts in stimulation. A rat, for instance, may be placed in the maze in slightly varying fashions, thus receiving different tactual, proprioceptive, and visual cues; and he may therefore make a response at the first choice point different from that which he had made there previously. This in turn changes partially the stimulus pattern operating at the next choice point, and so a

different response may be made there also. Similarly, the human *S* may radically shift his position, thus changing from one trial to the next his whole pattern of proprioceptive cues. In addition, there may be differences in the external stimuli present on different trials: changes in lighting or olfactory cues, noises outside the laboratory, motions by the experimenter, or any one of many other accidental stimuli. Occasionally therefore, the stimulus pattern is quite different from that present the last time the subject was at this choice point. Consequently, at this point a different response may be made to this different stimulus pattern; and the prediction (based upon the last behavior to a different stimulus pattern at the choice point) will be incorrect.

This does not mean necessarily that the principle of postremity is not completely valid. It may be. The principle states what response is to be expected when the same situation is re-encountered. However, for practical reasons, the principle is used to predict behavior when only a part of the situation is re-encountered (namely: the 'same' choice point). This then is the first reason why predictions based upon postremity, though remarkably accurate, are not perfect; and this is also the first reason why the *S* does not repeat forever the same response to any given choice point: The *S* is not the same *S*, and the choice point is not the same choice point from trial to trial.

B. Elimination of errors.

So far we have been concerned only with why, if the principle of postremity has a consistent effect, there are variations in responses. Going beyond that problem is the question of why errors to a great extent are eliminated.

The factor just discussed is a part of the answer to this question—changes in the internal and external stimulus patterns bring variations in response at the 'same' choice point. However, if this were the only factor operating and if these changes in the stimulus patterns occurred in a perfectly random fashion, errors would not be eliminated; for there would be (within the limits of chance) as many changes in stimulus pattern at times when the right response had been made as when the wrong response had occurred. Thus there would be, on the average, as large a proportion of shifts from making the correct response to making an error as from making an error to making the correct response at a choice point.[8] So the animal would not make the same error always; but neither would this error be eliminated.

If, on the other hand, these changes in the situation do not occur in a perfectly random fashion, errors could tend to be eliminated. It seems highly probable that this is the case, that there are indeed greater variations in the stimulus pattern when errors, than when correct responses, are made. It seems probable, for example, that preceding and accompanying errors

[8] More precisely, the ratio of (a) number of shifts (from one trial to the next) from 'right' to 'wrong' responses to (b) number of 'rights' would be equal on the average for all trials to the ratio of (a) number of shifts from errors to 'right' responses to (b) number of errors.

there is greater uncertainty and anxiety aroused by the stimuli present at those choice points and more emotion as the result of making errors (being blocked); hence, as these emotional response patterns characteristically involve varying tensions, there are greater changes in stimulation at those times. Similarly, it has been observed that many human beings tend to withdraw from a situation in which they are not being 'successful.' This would result in fluctuations of attention—i.e., involve many of the stimulus patterns being present for the *S* sometimes and not being present other times, and also would result in introduction of new stimuli by the *S*'s talking to himself about various changing matters. Thus, there would be, for this reason too, more variability of stimuli when errors are made than when the correct responses are made. It follows from these considerations that correct responses to choice points would have a greater tendency to be repeated on subsequent trials (because those stimulus patterns are more nearly constant from trial to trial); and the errors would tend less strongly to be repeated (because the stimulus patterns are more variable).

In addition to this, there is a second factor which, if the principle of postremity is valid, would result in the elimination of errors. The factor is this: Following a correct response to a choice point, usually no stimulus is given eliciting a different response at that choice point; whereas, in sharp contrast to this, following an incorrect response, frequently a stimulus *is* given which causes the animal to make some other response while many of the choice point stimuli are present. Thus at a choice point where the correct response was made on the last trial, the correct response usually is the postreme response to all stimuli present at that choice point. But where an error was made on the last trial, the error—although it is the postreme response to some of the stimuli—often is *not* the postreme response to many of the other stimuli present at that choice point.

For example, in a multiple-T maze, when the correct response is made at a choice point, there is usually no retracing; the subject goes on to the the next. Therefore, the last response that was made to all stimuli at that point was to go down the correct alley.

In contrast to this, when an error is made at a choice point the animal is faced by a wall or some sort of block, in response to which he comes back out of the blind and then goes down the correct alley. This is not, of course, identical to making the correct response to that choice point stimulus pattern which just elicited the wrong response. (And therefore, 'wrong' is predicted as the response which will be made the next time the animal arrives at that choice point.) Nonetheless, the stimulus-pattern-response-pattern of going down the correct alley after coming out of the blind has some components identical to that of going down the correct alley without going into the blind (e.g., sight of correct alley accompanies running down right alley in either case); and so even on those trials when the wrong response is made, some of the stimuli become cues for making the correct response. Also, many of the stimuli that accompanied and elicited the wrong response are present when the incompatible response later is made of turning around and coming out of the blind; and so these stimuli are detached from the response of continuing down the blind. Since these stimuli are then cues for coming

out of the blind and no longer cues for continuing to go down it, they indirectly are cues for going down the correct alley (if there are but two ways to go at a choice point).

In spite of this detachment of some of the choice point stimuli from the wrong response, as was previously mentioned there still are many stimuli at this choice point which are cues for making the wrong response (e.g., the stimuli from whatever responses were made at preceding points, of approaching this point, of the sight of the blind and other alley extending at certain angles from the approach point, etc.) Not until the animal approaches this point, and without first starting down the wrong alley, makes the correct response, can all stimuli of that choice point become cues for the correct response.

Thus it is seen that the choice point situation which has elicited a *wrong* response *becomes ambiguous* when the animal comes back out of the blind and goes down the correct alley—some of the stimuli are still cues for the wrong response but others now have become cues for the right response; whereas the choice point situation which last elicited a *correct* response usually is *not ambiguous*—all of the stimuli remain cues for the right response. Hence, in accordance with the present theory, right responses will tend more strongly to be repeated than will wrong.

Another example of how this second factor may operate toward facilitating elimination of errors is found in the consideration of behavior to the punch-board maze. The factor, it will be recalled, is this: Following a response which is an error, in many experiments a stimulus is given which may cause the animal to make some other response in the presence of much of the stimulus compound which had elicited the error. Now, in the learning of the punch-board maze, *S* may insert his stylus into the wrong hole of a pair, e.g., pair 8. A stimulus then is presented (the sound of the buzzer in our experiment, *E* saying 'wrong' in others, etc.) which may cause *S* to withdraw his stylus and make additional responses on that trial to that choice point. These responses may involve, for instance, the *S*'s saying to himself "Eight, not this hole, that one is right," or even making incipient or possibly overt movements toward the other correct hole while looking at the pair. If this has been the case, the next time the *S* comes to that pair of holes, he will say "Eight" (as instructed), and it is expected he will start toward the incorrect hole, draw his stylus back, say "No, not this hole; that one is right," and then make the previously established conditioned response to the stimulus "that-is-right,", i.e., the response of inserting his stylus into the hole at which he just looked and said "That is right." Thus another error may be eliminated. This again is in accordance with the principle of postremity: The *S*, when he previously approached that pair, made an error. This response was then followed by presentation of a stimulus which in this case caused *S* to make some other response to a large part of the stimulus pattern to which he had made the error (such parts would be the stimulus of the sight of the pair of holes, saying the number of the pair, etc.). Thus this 'other response' becomes cued to those parts of the stimulus pattern and the error is detached from them. The next time *S* came to the pair, he made the response which was his postreme response to most of the stimulus pattern.

This postreme response, it should be noted, is a response different from that of inserting the stylus into the wrong hole; and therefore the wrong response was not made on this trial.

C. Imperfect predictions.

There is one more point which should be noted in connection with the above. Predictions are based upon the last response *recorded;* whereas behavior (according to the principle of postremity) depends upon the last response *made* to a situation. Now if the response to the stimulus pattern of a choice point is correct, the last-recorded response is likely to be the last-made response (because, as previously suggested, the subsequent stimuli indicating that the response is correct usually do not bring additional responses to that choice point differing from that recorded); but if the response is incorrect, the stimuli following that response which indicate its 'incorrectness' often do bring other added responses on that trial to many of the stimuli of that pair, added responses such as movement away from the incorrect hole and toward the correct hole. These responses (which are now the postreme response to some of the stimuli of that choice point situation) are not recorded, however; and they are different from and incompatible with the one that was recorded. Therefore, since the last-made response sometimes does not agree with the last-recorded response (especially in the case of errors), and since predictions are based upon the last-recorded response while future behavior is a function of the last-made response (according to the postremity hypothesis), the predictions will fail sometimes to agree with the behavior even though the theory is sound.

Another main reason, previously mentioned, why predictions fail to have complete accuracy is that the stimulus pattern was changed from that of the preceding trial. As these two factors constitute a part of the conditions assumed to be fulfilled when one uses the principle of postremity in the way followed in this paper, it is imperative that the particular stimulus pattern present now and the particular response last made to these stimuli be noted with extreme care if one is to test the present theory. If they were taken more completely into account, predictions should have an accuracy still higher than that we found—providing the theory including the principle of postremity is valid.

D. Postremity and frequency.

The final problems we shall consider are why the postreme response was more apt to be repeated when it had been also the more frequent response than when it had been the less frequent; and why the postreme response was

repeated in many instances even though it had been the less frequent response. One interpretation in accordance with the principle of postremity is based in part upon the two following points.

1. For a certain choice point, if the response recorded as the postreme response of trial X was also the more frequent response, this indicates: (a) the stimulus pattern on X was very similar to ones that had occurred at that choice point frequently in the past (for if it were not, a response different from that made on most preceding trials would have occurred on X and also responses different from that which did occur most frequently on the preceding trials would have had to occur on those previous trials), and (b) the response recorded on preceding trials as the postreme response to that choice point generally was the response actually last made at that point.

2. If the response recorded as the postreme response of trial X′ was the response less frequently made in the past, this may indicate: either (a) the stimulus pattern present on X′ was radically different from ones which had most frequently appeared previously at that choice point; or else that (b) the response recorded for trial X′ − 1 was not actually the postreme reponse made at the choice point in question on that trial (trial X′ − 1), and this had not happened often before.

If 2-a is at least as common as 2-b, and if radically different stimulus patterns are apt to be transitory changes in the fundamental stimulus patterns—i.e., if the nature of the future in respect to stimulus patterns is apt to be more like the most common part of the past than like any other one segment of the past, it follows most directly from 1-a and 2-a that in general the stimulus pattern present on trial X will be expected to recur more often on the following trial than will the pattern of trial X′. Because of this, and because recurrence of the stimulus pattern which accompanied a response is essential for that response to be evoked again, those responses indicating that certain similar stimulus patterns have frequently appeared in the past theoretically should be the ones made on subsequent trials more often than those responses which may indicate the presence of stimuli which have but rarely occurred previously. On the basis of this reasoning, we should expect (as was found) that the response of a certain trial would be more apt to be repeated on the next trial when that response was also the one which had been the more frequent than when it had been the less frequent response.

It should be noted that this interpretation of why our predictions from postremity were more accurate when the response of the preceding trial was also the more frequent does *not* attribute this phenomenon to postremity perhaps operating more efficaciously at some times than at others, nor does it attribute the phenomenon to *frequency of the response per se;* but rather it is held as being due fundamentally to the relative frequency with which *certain stimulus patterns* have occurred in the past and therefore may be expected to recur in the future—a matter of which past frequency of the response is taken as an indirect indication. Guthrie has offered another interpretation which also may be applied to the present problem, one which

is fundamentally in terms of the frequency with which a response has occurred.[9]

We turn now to an interpretation of the final problem stated above. When the occurrence on trial X′ of the response made less frequently on preceding trials is due to stimuli having been applied on trial X′ − 1 which caused *S* to make an additional response at that point different from the one recorded as his postreme response for trial X′ − 1 (i.e., when the response on trial X′ is due probably to factor 2-b rather than 2-a), we should expect, if the present theory of postremity is valid, that for those times the response on trial X′ *would* be the one to occur on X′ + 1—under these conditions, the fact that frequency is opposed is irrelevant. Further, if such additional responses occur appreciably often, we should expect the number of times predictions based upon this hypothecated factor (2-b) and upon postremity (though frequency is opposed) accorded with subsequent behavior to be at least statistically significant. As was pointed out in part B of the preceding section, this was found.[10]

Summary and Conclusions

To determine the accuracy of postremity as a basis for prediction of behavior in the maze situation, the behavior of 57 *S*s was studied, two mazes of radically different type (the raised relief finger maze and the punch-board maze) and two different procedures on the last maze being used, one group of *S*s having to push the stylus in the holes, the other having to screw it in.

It was found that prediction of responses based upon postremity correspond remarkably closely to future observed responses—not only for all groups but for all individuals, the predictions being correct for no individual less than 73 percent of the time and for most individuals more than 85 percent of the time. For each individual, the accuracy was significant considerably beyond the 0.1 percent level of confidence.

High accuracy of predictions based upon postremity also was found even for those times on which it predicted the response opposed to the one most frequently made in the past. For all groups and

[9] This interpretation is that the stimulus pattern changes somewhat from trial to trial and so with added occurrences of the response more stimuli become cues for the response in question. Therefore on future trials, more of the stimuli which may occur have been established as CS's for the response. From this and from the hypothesis that occurrence of a response is positively correlated with the proportion of stimuli present (at the particular time for which prediction is being made) that is composed of CS's for the response, it follows that the more frequently a response has been made the more certain it is to re-appear on the next trial (cf. 5, pp. 99–101).

[10] The number of times postremity was accurate under these conditions and the statistical significance of these results should not be imputed to factor 2-b only; for even if the less frequent response of trial X′ was due to factor 2-a and if this stimulus pattern, though rare in the past, was still present on the next trial (which would at least occasionally be the case in our opinion), then for those trials also the same response would occur on X′ + 1 as on X′, and this also would contribute to the accuracy of postremity as we used it.

most individuals, the accuracy of these predictions exceeded 50 percent by an amount which was significantly greater than would occur by chance.

Predictions based upon frequency did not correspond well with future observed responses *unless these predictions were in accord with those of postremity*. In fact, when frequency and postremity were at variance, predictions based on frequency were less accurate than would be ones based upon chance. For one *S* only were they correct as much as 50 percent of the time. Statistical analyses showed that for the majority of the individuals and for all groups, the accuracy of these predictions from frequency was significantly less than chance and also significantly less than that of those based on postremity.

Suggestions were offered as to the conditions under which variations would occur in the responses made by the same *S* at the 'same' choice point, and under which errors would be eliminated if the principle of postremity were perfectly valid. Interpretations were offered also as to why predictions were imperfect, why the response recorded for the preceding trial was more apt to be repeated when it had also been the more frequent response than when it had been the less frequent, and why it was repeated in many instances even though it had been the less frequent response.

(Manuscript received August 27, 1947)

References

1. Brown, T. *The philosophy of the human mind.* London: Hallowell, Glazier, Masters and Smith, 1846.
2. Granneberg, A. O. *The factor of tuition in maze learning.* M. S. Thesis, Univ. Washington Library, 1937.
3. Guthrie, E. R. Conditioning as a principle of learning. *Psychol. Rev.*, 1930, 37, 412–428.
4. Guthrie, E. R. Reward and punishment. *Psychol. Rev.*, 1934, 41, 450–460.
5. Guthrie, E. R. *The psychology of learning.* New York: Harper, 1935.
6. Guthrie, E. R. Tolman on associative learning. *Psychol. Rev.*, 1937, 44, 525–528.
7. Guthrie, E. R. Conditioning: a theory of learning in terms of stimulus, response, and association. in: *Forty-first yearbook of the national society for the study of education, Part II.* Bloomington: Public School Publishing Co., 1942.
8. Guthrie, E. R. The principle of associative learning. In: *Philosopical essays in honor of Edgar Arthur Singer, Jr.*, Univ. Penna., Press, 1942.
9. Guthrie, E. R., & Horton, G. P. A study of the cat in the puzzle box. *Psychol. Bull.*, 1937, 34, 774.
10. Guthrie, E. R., & Horton, G. P. *Cats in a puzzle box.* New York: Rinehart, 1946.
11. Horton, G. P., & Guthrie, E. R. A further study of the cat in the puzzle box. *Psychol. Bull.*, 1938, 36, 521.
12. Hudson, B. B. One-trial learning in rats. *Psychol. Bull.*, 1939, 36, 643.
13. Hunter, W. S. Learning: IV: Experimental studies of learning. In: C. Murchison, ved., *A handbook of general experimental psychology*, Worcester: Clark Univ. Press, 1934, pp. 497–570.
14. Jersild, A. Primacy, recency, frequency, and vividness. *J. exp. Psychol.*, 1929, 12, 58–70.

15. McGeoch, J. A. *The psychology of human learning: an introduction.* New York: Longmans, Green, and Co., 1942.
16. Muenzinger, K. G., & Vine, D. O. Motivation in learning: IX. The effect of interposed obstacles in human learning. *J. exp. Psychol.*, 1941, **29**, 67–74.
17. Peterson, J. Frequency and recency in maze learning by white rats. *J. Anim. Behav.*, 1917, **7**, 338–364.
18. Peterson, J. Learning when frequency and recency factors are negative. *J. exp. Psychol.*, 1922, **5** 270–300.
19. Spence, K. W. The reliability of the maze and methods of its determination. *Comp. Psychol., Monogr.*, 1932, No. 40.
20. Treloar, A. E. *Elements of statistical reasoning.* New York: Wiley and Sons, 1939.
21. Warden, C. J., & Cumming, S. B. Primacy and recency factors in animal motor learning. *J. genet. Psychol.*, 1929, **36**, 240–256.
22. Watson, J. B. *Behavior, an introduction to comparative psychology.* New York: Holt, 1914.
23. Yoshioka, J. G. Pattern vs. frequency and recency factors in maze learning. *J. genet. Psychol.*, 1928, **35**, 193–200.

ON THE PROBLEM OF 'REINFORCEMENT' IN CONDITIONING THE AUTOKINETIC PHENOMENON

BY ERNEST A. HAGGARD

University of Chicago

AND

RACHEL BABIN

Radcliffe College

The concept of reinforcement is fundamental to learning theory. But because of the existence of widely divergent theoretical systems, the supposed nature and role of reinforcement differ considerably for different writers. To some the law of effect (or some modification of it) is of primary importance as an explanatory principle; by others the law of effect is held to be of negligible significance in the learning process. Another source of disagreement in this field is the fact that a given set of data frequently can be rationalized in terms of more than one theoretical framework. As a consequence, there is perhaps more disagreement than agreement as to how learning is to be 'explained.'

The present study was designed to investigate some aspects of the role of reinforcement in the learning process. There is, however, one difficulty that exists because of the very nature of the autokinetic phenomenon. The conditioned response in most learning experiments is almost always sufficiently manifest to enable the *E* to measure and record various characteristics of the particular behavior under investigation. Thus, the obviousness of such responses as a knee-jerk or the path taken by a rat at a choice point in a maze gives the *E* confidence that he is observing something. But in working with the autokinetic phenomenon, the *E* is unable to make any such direct observations. Rather, it is necessary to rely on the verbal report of the *S*, and to assume that a sufficiently high correspondence exists between the *S*'s subjective perception and his report of that perception. Consequently, such expressions as 'apparent movement of the light' or movement which is 'seen,' 'perceived,' 'reported,' etc., will be considered synonymous.

In a previous study on the autokinetic phenomenon (**18**), *S*s were conditioned to see the light move more frequently, farther, etc., in a leftward direction. Differences in the mental set of the *S*s, and whether they assumed an active or a passive role in the learning

situation, were influential in determining the effectiveness of this conditioning. Reward reinforcement was present for all *S*s.

In the course of this earlier investigation, three questions emerged, namely: (a) could comparative statements be made concerning the relative effectiveness of reward reinforcement and motor reinforcement (active participation) in conditioning the autokinetic phenomenon, (b) if motor reinforcement is of major importance, does the 'pattern' of such reinforcement play a substantial role in the conditioning procedure, and (c) whether the direction in which the light was conditioned to move was important, and if so, was this factor related to eye dominance.

The present paper will report an attempt to investigate these questions, and will consider certain theoretical implications of the results of this study.

The Experiment

Subjects

The 18 *S*s in this experiment were volunteers from psychology classes in Harvard and Radcliffe colleges. All of the *S*s were unaware both of the purpose of the study and of the fact that the position of the light was fixed during all five sessions. When, at the close of the experiment, they were told the nature of the study, some were surprised, some amused, some indignant, and some stated that they could not believe that the stimulus had remained stationary, so definite had been their perception of the (apparent) movement.

Apparatus

The apparatus consisted of a 7-watt bulb connected to a throw switch and was mounted in a small light-tight tin can. One side of the can was pierced with a pin-point hole to provide the stimulus. The shelf on which the light rested was three ft. above the floor and was calibrated by cardboard tabs placed at two in. intervals. The tabs served to provide the *S* with a scale which he could retain subjectively when the room was darkened, in order to estimate the distance that the light 'moved' during any given trial. Ostensibly, these tabs also seemed to provide the *E* with a means of determining the extent of the movement of the stimulus during each trial. The *S* sat 10 ft. away from the stimulus in an upright chair. Except for the pin-point light, the room was completely blacked out after the experiment started.

Experimental variables

I. *Motor reinforcement*

The 'Movement With' group (MW).—The six *S*s in this group were provided with a pencil and pad of paper (11 × 8½ in.) interleaved with carbon paper, and were asked to draw a heavy arrow the same length and pointed in the same direction as the perceived movement of the light during each trial. The *S* drew the arrow as the light moved. If no movement was observed, the *S* made a small circle in the center of the page for that trial. Movements of the light other than horizontal were ignored. At the end of each trial, the *S* also said whether the light moved, and if so, which direction, and how far.

The 'Movement Against' group (MA).—The six *S*s in this group experienced the same general conditions as did the MW *S*s, except that they were asked to draw arrows in the direction opposite to the apparent movement of the light. For example, if the light seemed to move 2½ in. to the left, they were to draw an arrow 2½ in. to the right, and vice versa. Thus all conditions except the direction of the motor response were comparable to those of the MW *S*s.

The 'No Movement' group (NM).—The six *Ss* in this group sat passively during all the trials, and were only required to say whether the light moved at all during each trial, and if so, the direction and distance of the (horizontal) movement.

II. *Reward reinforcement*

The 'Reward' group (R).—In the earlier experiment (18), all *Ss* were 'rewarded' for reporting movements to the left and 'punished' for movements reported to the right. Two modifications were introduced in the present study: (a) the rewarded direction, and (b) the presence or absence of reward reinforcement. Thus half, i.e. nine, of the *Ss* were rewarded for reporting movements to the right and punished for the movements to the left. The 'reward' used here included both personal or social as well as monetary aspects. That is to day ,the *E* said 'Correct' and credited the *S* with five cents, or said 'Wrong' and required the *S* to forefeit five cents, depending on whether the light was reported as having moved to the right or to the left. Also, after the first two sessions, the *E* told each *S* that some of the other *Ss* were doing better, and hence were earning more money than he (or she) was.

It was necessary, however, to complicate the system of rewards and punishments so that the *Ss* would not become aware of which responses were considered correct or wrong. This was done by randomizing the presence or absence of reward reinforcement, so that only half of the movements to the right were rewarded, half of the movements to the left were punished, and the no-movement responses were alternately rewarded and punished. On the non-reinforced trials, the *S* was told, "That trial didn't count." The order of reinforcements was as follows: oxxoxooxoxoxxox, etc., where 'o' signifies the non-reinforced and 'x' the reinforced trials. This method was apparently successful, since none of the *Ss* were certain of the system of rewards and punishments by the end of the experiment.

The 'No Reward' group (NR).—The nine *Ss* in this group were neither rewarded nor punished, and never knew whether their reported perceptions were 'correct' or 'wrong,' since the *E* remained noncommital.

III. *Eye dominance*

The 'Right-Eyed' (RE) and 'Left-Eyed' (LE) groups.—Because of the possibility that factors associated with eye dominance may bear a close relation to cerebral dominance, and hence perhaps to the autokinetic phenomenon, the *Ss* were divided according to whether they were Right Eyed (RE) or Left Eyed (LE). The eye dominance of each *S* was determined by a simple card-hole test. The *S* was required to look through a small hole in the center of an $8\frac{1}{2} \times 11$ in. piece of cardboard in order to distinguish between a series of open and closed circles held five ft. away. Since the size of the circles ranged from $\frac{1}{4}$ to $\frac{3}{4}$ in., the *S* was led to believe that he was being given a test of visual acuity. Although the chart could be seen with only one eye, no directions were given as to which eye was to be used. The eye which was used to look through the hole was called dominant. According to this test, nine RE and nine LE *Ss* were selected.

In order to minimize the possibility of confounding the factors involved in eye dominance with those involved in motor and reward reinforcement, the RE and LE *Ss* were assigned at random to the other conditions, with the exception that not more than two of a kind (i.e. either RE or LE *Ss*) should be placed in any given sub-group. Consequently, there were three RE and three LE *Ss* in the Movement With group and four RE and five LE *Ss* in the Reward group. Thus it is relatively safe to assume that the factor of eyedness was not inadvertantly confounded with the arrangement of *Ss* and treatments in the experimental design.

Procedure

At the beginning of each session, the *S* was seated in an upright chair 10 ft. away from the stimulus (the pin-point light). After the *S* read the set of instructions at the beginning of Session I, and any questions as to the procedure were cleared up, the room light was turned off for 30 sec. to allow the *S* to become accustomed to the darkness before the first trial. A trial consisted of the presentation of the stimulus light for 15 sec.; the interval between trials was 30–35 sec. When the stimulus light was turned off at the end of each trial, all *Ss* reported whether the light moved, and if so, the distance and direction of this movement. These verbal

reports were recorded by the *E*. Each session was comprised of 30 trials. The room light was not turned on until the end of the session. Each *S* received only one type of treatment during Sessions 1–5.

Instructions

Sessions 1 to 5.—Because of variations in the experimental procedure, six forms of typewritten instructions were prepared. At the beginning of the first session, the *S* was told about the importance of developing tests for the perception of small movements, in order to select airplane spotters, etc. This was done to increase motivation and to disguise the real purpose of the experiment. After this introduction, the *S* was handed the appropriate set of instructions. The form given to the *S*s in the Movement With—Reward group was as follows:

> During this experiment, the room will be completely dark. After you have become used to the darkness, you will see a point of light appear in front of you, about 10 ft. away. The light will stay on for 15 sec. It may move to the left, to the right, or it may not move at all. In the trials when the light does move, follow the light *as it moves* with your pencil on the pad of paper in front of you. For example, if you see the light move 4 in. to the left (or to the right) draw an arrow 4 in. long pointing to the left (or to the right). Disregard all movements except the horizontal ones. If the light does not move at all during a particular trial, draw a small circle in the center of the page.
>
> Put the number of the trial before each line or by each circle, and when you draw the line remember to bear down hard enough so that the impression will carry through at least 3 carbon sheets.
>
> When the light in front of you is turned off, tell me whether it moved, and if it did move, report as accurately as you can the direction and distance of the movement. For example, after each trial say, 'Left, 4 inches,' or 'Right, 2½ inches,' or 'No movement.'
>
> There will be 30 trials in each session. Remember to report only the horizontal movements, and that the light may not move during every trial. Also remember that only about half of the trials will count, but you will not know ahead of time whether the light will move or if the trial will count.
>
> After each trial you will be told whether you were correct or wrong, or if that particular trial didn't count. Remember that every time you are correct you will earn 5 cents, but whenever you are incorrect you will have to forfeit 5 cents. So be alert and don't be caught off guard.
>
> Do you have any questions?

Session 5b.—At the end of trial 30 of Session 5, all rewards, punishments, and knowledge of results were suspended and the *S*s were asked to go through an additional 10 trials. During these trials, the *S*s were asked to report the direction and distance of the perceived movements, and to *tell how certain or confident* they were (along a 1–5 scale) that their perceptions were correct. A score of five was to be assigned if they were very certain of their perception of the movement, and a score of one if they were very uncertain. In making these estimations they were asked to pay more attention to the direction of the movement than to its magnitude.

Measures

Index of conditioning.—This measure is the sum of the percent change in the direction and in the orally reported distance of the movements seen to the right (the rewarded direction) between Sessions 1 and 5. That is to say, this measure is the algebraic sum of (a) the increase in the proportion of all movements seen to the right added to (b) the increase in the magnitude of the movements seen to the right between the first and fifth sessions. Thus the larger the positive score on this index, the more 'conditioning' is assumed to have taken place. A negative score indicates a decrement in the frequency and magnitude of movements seen to the right between Sessions 1 and 5, and may be due to the nature of the experimental treatment or some undetermined structural-functional properties of the nervous system or both.

Frequency of movement.—This rubric includes three separate measures, namely: (a) the average percent of movements seen to the right, (b) the average percent of movements seen to the left, and (c) the average percent of trials during which no movement was reported.

Magnitude of movement.—The magnitude of the movements, kept separate for those to the left and to the right, were estimated in terms of inches, and were based on the *Ss*' verbal report.

Degree of confidence.—These measures were obtained by computing the average of the *S*'s reported confidence of certainty of his perceptions of (a) movements to the right, (b) movements to the left, and (c) no movements during the 10 trials following Session 5. (See *Procedure, Session 5b.*)

Results [1]

Variable I: motor reinforcement

It will be recalled that the Index of Conditioning used in this study is a compound measure, and that a positive score signifies an increase in both the frequency and magnitude of the apparent movements seen to the right between Sessions 1 and 5. From Table

TABLE I

Mean Values for Subjects on the Various Measures in Terms of the Three Major Variables

Measures	Session	Variable I Motor Reinforcement			Variable II Reward Reinforcement		Variable III Eye Dominance	
		Mvt. With	Mvt. Against	No Mvt.	Reward	No Reward	Left Eyed	Right Eyed
I. Index of conditioning	1–5	38.3	−32.1	−69.0	−3.7	−38.2	6.1	−47.9
II. Reported frequency of perceived movements	5							
a. To the right		46%	36%	30%	42%	34%	43%	30%
b. To the left		42%	33%	38%	37%	34%	31%	40%
c. No movement		12%	30%	32%	20%	31%	28%	30%
III. Reported magnitude of perceived movements	5							
a. To the right		2.6 in.	1.8 in.	1.9 in.	2.8 in.	2.2 in.	2.0 in.	2.1 in.
b. To the left		2.6 in.	1.9 in.	1.3 in.	2.0 in.	2.5 in.	2.4 in.	2.1 in.
IV. The *Ss*' confidence in their perceptions	5-b							
a. To the right		4.0	3.0	3.2	3.4	3.4	3.5	3.4
b. To the left		3.0	3.3	2.7	2.7	3.2	2.8	3.2
c. No movement		3.5	3.4	3.6	3.5	3.5	3.2	3.8

I observe that differences exist between the groups receiving variations in the conditions of motor reinforcement. The Movement With (MW) group showed an average increase of 38.3 points on the

[1] Many of the results presented in this paper do not show clearly significant differences between groups of *Ss*, although they do indicate rather consistent trends. The results might have been more clear-cut if (a) the stimulus were presented for more than 15 sec., since according to one study (14, p. 115) the response did not appear for 50 percent of the *Ss* before about 13 sec. had elapsed, and (b) if 100 percent rather than 50 percent reward reinforcement had been used. But the use of 100 percent reinforcement would probably have resulted in the *S*'s discovery that movement to the right was always called 'Correct,' and vice versa, and this would have precluded their usefulness as *Ss* in this experiment.

Index of Conditioning, the Movement Against (MA) group a decrease of 32.1, and the No Movement (NM) group a decrease of 69.0.[2] The significance of the difference between the MW and MA groups falls at the .05 level ($t_{10} = 2.31$), the difference between the MA and NM groups falls at the .18 level ($t_{10} = 1.40$)[3], and the difference between the MW and NM groups falls beyond the .01 level ($t_{10} = 3.21$). Thus, it is indicated that congruent motor reinforcement, or reinforcement in a pattern which is not in conflict with the apparent movement of the stimulus, was the most effective single condition to serve as a facilitating state of affairs in conditioning the autokinetic phenomenon in this experiment. Furthermore, the presence of motor reinforcement was more effective in facilitating learning than was a condition in which the *Ss* remained passive. This is evidenced by the fact that when the MW and MA groups were combined to form a Movement group, and compared with the NM *Ss*, a difference at the .05 level of significance ($t_{16} = 2.37$) obtained.

Certain interesting trends also appear when we consider the relative frequency of the light's movement during Session 5, whether it appeared to move to the right, to the left, or did not seem to move at all. As would be expected from the results just presented, the MW group saw it move to the right most frequently (46 percent of the trials), and was followed by the MA group (36 percent), whereas the NM *Ss* perceived the fewest number of movements in this direction (30 percent). Similarly, the MW group perceived the greatest frequency of movements to the left during Session 5 (42 percent), and was followed by the NM group (38 percent) and finally the MA *Ss* (33 percent). This necessarily means that *Ss* in the MW group experienced the fewest trials during which the light did not appear to move (12 percent), and the *Ss* in the MA and NM groups reported no movement during 30 and 32 percent of the trials respectively. In view of these findings, it appears that both the fact and the direction of motor reinforcement were of importance, since both the MW and MA groups showed a greater frequency of movements to the right than to the left, whereas the NM *Ss* reported more movements to the left than to the right.

[2] The reason for this decrement may be a predominant tendency for the *Ss* in this experiment to see the light drift to the left more than to the right, since (a) no such negative trend occurred in the study (18) in which the autokinetic was conditioned to the left, (b) the control sub-group, the NM-NR *Ss*, reported the greatest shift to the left (a mean score of − 97.4), (c) the rank order correlation for all *Ss* between their reported frequency of preceived movement to the left and their confidence in the correctness of these perceptions was .87, whereas the same value for movements seen to the right was .06, and because (d) certain trends appeared between eye, or cerebral, dominance and the extent to which conditioning occurred in this experiment.

[3] If a change in only the direction of the movements were considered (i.e., without respect to their magnitude), the difference would lie at the .02 level ($t_{10} = 2.99$). The subscript of the *t*-values refers to the appropriate number of degrees of freedom.

The importance of the factors involved in motor reinforcement was also manifest with regard to the distance that the light seemed to move during the final session.[4] That is to say, the *Ss* in the MW group saw the light move farther to the right (2.6 in.) than did the MA (1.8 in.) or the NM (1.9 in.) groups. Similarly, the MW *Ss* saw it move farther to the left (2.6 in.) than did either the MA (1.9 in.) or the NM (1.3 in.) groups. The fact that the MA and NM groups saw it move about the same distance to the right may be accounted for if we assume that the Movement Against condition both facilitated (as a result of the active participation) and inhibited the conditioning of the autokinetic phenomenon (as a result of incongruent motor reinforcement), so that the final effect was about the same as that for no motor reinforcement.

The results pertaining to the *Ss*' confidence or certainty of their judgments of distance and direction of the movement (Session 5b) also fell in line with the above findings. With regard to confidence in movements perceived to the right, the *Ss* in the MW group were on the average most certain of their perceptions (4.0), whereas the average score of the MA *Ss* was 3.0 and that of the NM *Ss* was 3.2. The difference between the MW and MA groups on this measure was significant at the .05 level ($t_{10} = 2.25$). As for the other conditions, *Ss* in the MA group were more confident when they saw the light drift to the left (3.3) than were either the MW (3.0) or the NM (2.7) groups. The fact that the MW *Ss* were most confident of their perceptions of movements to the right, and the MA *Ss* of movements to the left, implies that the factors involved in the direction of motor reinforcement were also effective in determining the *Ss*' confidence in their judgments of the apparent movement of the light.

Variable II: reward reinforcement

Reward reinforcement was not so effective as congruent motor reinforcement in conditioning the autokinetic phenomenon in this experiment. However, that is not to say that 'reward' played a role of little or no importance in this study, because it is necessary that the *S* be provided with some means (even though subliminal) of making discriminations between a 'correct' and an 'incorrect' response.

Although *Ss* in the Reward (R) group showed a slight decrement (−3.7) on the Index of Conditioning, the No Reward (NR) *Ss* dis-

[4] For the 12 *Ss* in the Movement group, the rank-order correlation between the measured length of the drawn lines and the estimated magnitude of the perceived movements was .64, between the length of the drawn lines and the frequency of the perceived movements was .38, and for all 18 *Ss*, this correlation between the frequency of the perceived movements and the estimated magnitude of the movements was .56.

played an even larger decrement (−38.2). The difference between these groups was not statistically significant, since the *P*-value falls at the .23 level ($t_{16} = 1.21$). But it is possible that if there had not been such great variability within these groups as a result of variations in the conditions of motor reinforcement and eye dominance, this difference might have approached or even reached significance.

When these reward groups are compared in terms of the total frequency of movements seen during Session 5, it is found that the *Ss* in the R group saw movements to the right during 42 percent of the trials, whereas the NR *Ss* reported movements in this direction during only 34 percent of the trials ($t_{16} = 1.18$; $P = .28$). During this session, the R group also saw slightly more movements to the left, and the NR *Ss* reported seeing no movement during 11 percent more of the trials (31 vs. 20 percent) than did the R *Ss*. This difference, however, again did not approach significance ($t_{16} = 1.18$; $P = .28$).

A similar trend with regard to the magnitude of the movements appeared during Session 5, in that the R *Ss* reported larger movements

TABLE II

MEAN VALUES FOR THE SUBJECTS IN THE VARIOUS GROUPS AND SUB-GROUPS IN TERMS OF THE INDEX OF CONDITIONING

	Movement With	Movement Against	No Movement	Totals
Reward	60.6	−31.3	−40.6	− 3.7
No reward	16.0	−33.1	−97.4	−38.2
Totals	38.3	−32.1	−69.0	−20.9

to the right (2.8 in.) than did the NR *Ss* (2.2 in.), whereas the NR *Ss* saw larger movements to the left (2.5 in.) than did the R group (2.0 in.). Likewise, *Ss* in the NR group had a somewhat greater degree of confidence in their perceptions of movements to the left than did the R *Ss*, although there was little if any difference between these two groups with regard to their confidence in their perceptions of movements to the right and of no movement.

Finally, it is apparent from Table II that the factor of reward reinforcement operated in a manner similar to congruent motor reinforcement (MW), in that for the R *Ss* there was a greater degree of conditioning, and during Session 5, these *Ss* reported larger and more frequent movements of the light to the right. Conversely, the NR *Ss* were characterized by less effective conditioning, and reported a greater frequency of movements to the left during Session 5.

Variable III: eye dominance

A meaningful pattern of differences, and one which parallels the other findings, appeared between the RE and LE groups. For instance, the LE *Ss* reported a slight increment in the relative frequency and magnitude of movements to the right (the Index of Conditioning) during the experiment (6.1), whereas the RE *Ss* showed a decrement on this index of −47.9. The significance of this difference falls at the .08 level ($t_{16} = 1.88$). This is true in spite of the fact that the effects of motor and reward reinforcement were randomized.

With regard to the frequency of movements occurring during Session 5, the LE *Ss* reported movements to the right for 43 percent of the trials, as compared with 30 percent for the RE *Ss*. The significance of this difference lies at the .06 level ($t_{16} = 2.05$). Again, in connection with movement reported to the left, the RE *Ss* gave a somewhat higher value (40 percent) than did the LE *Ss* (31 percent), whereas the difference with regard to no movement was negligible. There was little if any discrepancy between these groups in terms of the magnitude of the reported movements.

The scores indicating the *Ss*' confidence in their perceptions again reflect differences between the RE and LE groups, with the former more certain of movements to the left (3.2) and no movement (3.8) than the latter (2.8 and 3.2 respectively). Both of these differences lie at the .22 level of significance. There was no appreciable difference in terms of movements seen to the right. Thus, from these findings it appears that whatever factors are associated with conditions of eye dominance (as defined by the card-hole test) served to influence the conditioning of the autokinetic phenomenon in this experiment.

Discussion

Various attempts have been made to 'explain' the autokinetic phenomenon.[5] Probably the most useful explanation, and the one which best accounts for the known facts, is in terms of some sort of 'streaming phenomenon'—even though investigators disagree on the nature, extensiveness, and locus of these streams. For some, they

[5] Although Skolnick (34) recently made out a case for the causal relation between eye movements and the autokinetic phenomenon, his evidence was not conclusive. For example, (a) he reported (after Exner) that when a small bright spot is placed in the center of a dimly luminous disk, the two are seen to move independently, and (b) with regard to the specific relation between eye movements and the autokinetic phenomenon, the illusion may appear prior to, or even in the absence of, eye movements. Furthermore, whereas he did show the existence of some relationship between eye movements (as in the case of nystagmus) and the autokinetic phenomenon, his data may also be interpreted in terms of the effect of some central state of affairs on peripheral events—namely that nystagmus may be an adjustive mechanism to maintain certain visual constants in the face of central fluctuations.

have appeared to be rather discrete, slowly moving swirls which pass over only a portion of the visual field (e.g., **9**, **10**), whereas others (e.g., **15**) have described streaming which covers the entire visual field. The figure-ground relationship between the pin-point light and the surrounding streams is assumed to be analogous to the apparent movement of the moon behind floating clouds (**16**).

Most writers have postulated that retinal processes underlie the observed streams, hence that the autokinetic phenomenon is determined by retinal events. However, in a previous paper (**18**), evidence was presented which seems to justify the assumption that their locus is cortical instead of retinal. These lines of evidence include the fact that the focus of the alpha rhythm is not fixed, but may pass several centimeters across the cortex (**2**), and that the migraine scotomas pass from the center to the periphery of the visual field, and are symmetrical for both eyes (**21**). These findings indicate that molar patterns of excitation pass across the cortex.[6] Furthermore, the autokinetic phenomenon is at least partially subject to voluntary control (e.g., **1**, **8**, **14**), suggestion (**33**), and conditioning (**18**). On the basis of such considerations, it appears that both the streaming and autokinetic phenomena are grounded in cortical events.

With regard to the variable of motor reinforcement, it will be recalled that the Movement with condition (congruent motor reinforcement) was the most effective single condition to modify the autokinetic phenomenon in this experiment. This is reflected in the over-all Index of Conditioning, in the frequency and magnitude of movements seen to the right during Session 5, and in the *S*s' reported confidence in their perceptions. The Movement Against condition (incongruent motor reinforcement) was less effective in these respects, and the No Movement condition (no motor reinforcement), where the *S*s remained passive, was the least effective.

Two sets of findings reported in the literature seem relevant to these results. One set pertains to experiments in which the fact of overt motor reinforcement plays a crucial role in the establishment of conditioned responses (CRs). Perhaps the clearest case in point is the group of so-called voluntary control experiments. In these studies, an essential factor in the conditioning procedure is that the *S* make a self-induced overt response in conjunction with the conditioned stimulus (CS). This voluntary response comes to serve as a sufficient condition to elicit the newly-acquired CR. Reward reinforcement, in the customary sense, is not present in these experi-

[6] Note also that the fruitfulness of postulating molar cortical 'field forces' in the study of perceptual problems has been demonstrated in the case of both apparent and real movement (**6**).

ments.[7] Examples of such CRs include the contraction of the retrahens muscle of the ear (**3**), pupillary (**19**) and vasomotor (**23**) changes, and blocking of the alpha rhythm or EEG (**20, 31, 32**).

A second relevant type of finding is concerned more specifically with the effect of the pattern of overt movement on the acquisition of the CR. A study by Beck (**5**) illustrates this. In his experiment, a tapping movement cycle of the hand served as the basis for conditioning the eyelid response to shock. He found that the organization of the movement cycle was a sufficient and necessary condition for the occurrence of conditioning.[8] Furthermore, the conditioned wink fell in step with the beat-stroke of the movement cycle, regardless of the temporal position of the shock, or 'UCS.' That is to say, "the position of the shock in the cycle had no effect upon any phase of the conditioning process here presented" (p. 395).[9] Thus the overt motor response was the essential reinforcing or facilitating factor in the establishment and maintainence of these CRs. But with specific reference to the temporal relation between the electric shock and the eye-wink (i.e. the traditional UCS and CS), the overt movement cycle interfered with, rather than facilitated, conditioning.

To summarize the discussion thus far, the following positions have been taken: (a) molar patterns of excitation passing across the cortex were postulated to underlie the autokinetic phenomenon, (b) overt movements made during the conditioning procedure may influence the acquisition of various CRs, and (c) the pattern of such overt movements may either facilitate or interfere with a particular CR.

In designing the present experiment two assumptions were made regarding the variable of motor reinforcement: first, that *S*s who assumed an active role in the learning situation would be more effectively conditioned than those who remained passive (cf. **17, 18**), and second, the nature or pattern of the motor reinforcement should be reflected in the relative efficiency of the conditioning under these two procedures. That is to say, it was hypothesized that molar patterns of excitation, one from the motor area and corresponding to the hand movement, and the other from the visual area and corresponding to the 'streams' associated with the autokinetic phenomenon, should somehow interact. Under such conditions, one might expect that the pattern from the visual area would be reinforced in the case of

[7] Note, for example, that in the case of the conditioned EEG, "the 'reward' demanded by the law of effect is not provided at a conscious level" (**32**, p. 207).

[8] Beck says that "organization of the movement cycle to a series of clicks is defined by the occurrence of one tapping movement when one click is omitted" (**5**, p. 396). In the absence of the organized movement cycle, conditioning failed to occur.

[9] It is possible, though, that the shock served as a sort of 'energizer,' as in the experiments by Finch and Culler (**11, 12**).

congruent motor reinforcement and interfered with in the case of incongruent motor reinforcement. In general, the findings of the present experiment conform to these assumptions, namely, that both the fact and pattern of motor reinforcement were influentual in determining the extent to which the autokinetic phenomenon was modified in this experiment.

The variable of reward reinforcement was included in this study in order to be able to make comparative statements concerning its relative effectiveness with motor reinforcement in modifying the autokinetic phenomenon. Two points are readily apparent from this comparison. One is that the pattern of the results for the Reward *Ss* is similar to that of the Movement With *Ss*, in that both groups reported larger and more frequent perceptions of movements of the light to the right during Session 5. The other point is that this trend is much less pronounced in the case of reward reinforcement, and the difference between the two groups is not statistically significant.

In evaluating the nature of 'reward' and its relation to the law of effect in particular experimental situations, it appears that two factors, or types of reward, are often operant. One type of 'reward' is essentially some form of information which enables the organism to discriminate between the 'correct' (or to-be-rewarded) and the 'incorrect' (or to-be-punished) responses. This seems to be primarily involved in the various cases of insight learning. The other type of 'reward' is usually defined in terms of the desirable gratification of some operant need, or the reduction of some undesirable inner tension system. But this latter type of 'reward' is dependent on, and essentially meaningful in terms of some, form of motivation which determines and directs the organism's performance. It is found most frequently in instrumental act learning situations.

In the present experiment, both types of 'reward' were apparently operant. But in the case of the former, it is necessary to assume that the *Ss* made discriminations below the threshold of awareness. This is required since none of the *Ss* were aware by the end of the experiment that actually the light had never moved at all, or that they were rewarded and punished in a pre-arranged systematic—although disguised—manner. There is ample evidence, however, to justify the position that subliminal discriminations occur (cf. **24**, **25**). The other contributing factor to the superiority of the scores of the Reward group may be accounted for by the function of an induced state of autism—the tendency of cognitive processes to move in the direction of need satisfaction, or tension reduction. The fact that 'unconscious' mental sets influence 'conscious' mental processes is

also well documented in various contexts (cf. **4**, **7**, **17**, **18**, **22**, **27**, **28**, **29**, **30**, **33**). Thus, it appears that reward reinforcement served, on a subliminal level, both (a) to increase the *Ss*' discriminations between the 'correct' and 'incorrect' responses, and (b) to modify their motivational pattern, and hence modify their response or performance pattern. Finally, there seems to be no reason to preclude the possibility that reward and motor reinforcement did not, somehow, interact to facilitate the conditioning of the autokinetic phenomenon in this study.

In returning to the apparent relationship between the streaming and autokinetic phenomena, several writers have made observations which indicate a further relationship between these phenomena and eyedness. More specifically, for a single eye, the streams generally have been observed to take a direction toward the center of the visual field (e.g., **9**, **10**), and to differ for the two eyes (e.g., **9**). Also, for most observers, when one eye is used the light has a predominant tendency to move toward the periphery of the visual field, and to differ for the two eyes (e.g., **1**, **8**, **13**, **16**). When both eyes are used, a rivalry generally exists, with one eye dominant, or there may be a shift of dominance, or some irregular resolution of the opposing movement patterns may occur. Gould (**13**) also supposed that "the character of the movements will depend upon whether one is right-eyed or left-eyed" (p. 536). (However, large individual differences exist in the extent and direction of the apparent movement under both monocular and binocular conditions.)

The eye dominance variable was included in the present experiment to investigate such relationships, and it was found that the Left Eyed *Ss* showed a larger score on the Index of Conditioning. Finally, whether eye dominance is innate or acquired, this finding agrees with others (e.g., **26**) that existent structural-functional properties of the nervous system determine to some extent the effectiveness of conditions external to the organism which facilitate the modification of behavior patterns, or learning.

Summary

This experiment was designed to investigate the following three questions: (a) whether comparative statements could be made concerning the relative effectiveness of reward and motor reinforcement in conditioning the autokinetic phenomenon to drift to the right, (b) whether the pattern of motor reinforcement played an important role in the conditioning procedure, and (c) whether the direction which the light was conditioned to move was important, and if so, whether this factor was related to eye dominance.

There were three experimental variables: (a) *Motor reinforcement:* Movement With (drawing arrows in the same direction as the apparent movement of the light), Movement Against (drawing arrows in the opposite direction), and No Movement (sitting passively); (b) *Reward reinforcement:* Reward (the presence of both social and monetary reward) and No Reward (their absence); and (c) *Eye dominance:* whether the *S*s were Right Eyed or Left Eyed by the card-hole test.

The measures used included an Index of Conditioning, the frequency, direction, and magnitude of all horizontal movements of the light, and the degree of the *S*s' confidence in their perceptions of the apparent movement of the light.

The results permit the following general conclusions:

1. Both the fact and the pattern of motor reinforcement played an important role in conditioning the autokinetic phenomenon in this experiment. The order of effectiveness of the conditions in this variable was: Movement With, Movement Against, and No Movement.

2. The reward reinforcement and eye dominance variables were less effective than motor reinforcement, with Reward and Left Eyedness being superior to the opposing conditions.

3. The difference in the relative effectiveness of the three variables in modifying the autokinetic phenomenon seems to be one of degree.

(Manuscript received July 7, 1947)

References

1. Adams, H. F. Autokinetic sensations. *Psychol. Monogr.*, 1912, **14**, No. 59, 1–45.
2. Adrian, E. D., & Yamagiwa, K. The origin of the Berger rhythm. *Brain*, 1935, **58**, 323–351.
3. Bair, J. H. Development of voluntary control. *Psychol. Rev.*, 1901, **8**, 474–510.
4. Bartlett, F. C. *Remembering.* Cambridge: The Univ. Press, 1932.
5. Beck, L. H. Conditioning and the coordination of movements. *J. gen Psychol.*, 1939, **20**, 375–397.
6. Brown, J. F., & Voth, A. C. The path of seen movement as a function of the vector-field. *Amer. J. Psychol.*, 1937, **49**, 543–563.
7. Carmichael, L., Hogan, H. P., & Walter, A. A. An experimental study of the effect of language on the reproduction of visually perceived form. *J. exp. Psychol.*, 1932, **15**, 73-86.
8. Carr, H. A. The autokinetic sensation. *Psychol. Rev.*, 1910, **17**, 42–75.
9. Edridge-Green, F. W. Visual phenomena connected with the yellow spot. *J. Physiol.*, 1910, **41**, 263–275.
10. Ferree, C. E. The streaming phenomenon. *Amer. J. Psychol.*, 1908, **19**, 484–503.
11. Finch, G., & Culler, E. Higher order conditioning with constant motivation. *Amer. J. Psychol.*, 1934, **46**, 596–602.
12. Finch, G., & Culler, E. Relation of forgetting to experimental extinction. *Amer. J. Psychol.*, 1935, **47**, 656–662.
13. Gould, G. M. A hitherto undescribed visual phenomenon. *Science*, 1903, **18**, 536–537.
14. Graybeil, A., & Clark, B. The autokinetic illusion and its significance in night flying. *J. Aviat. Med.*, 1945, **16**, 111–151.

15. GUILFORD, J. P. Autokinesis and the streaming phenomenon. *Amer. J. Psychol.*, 1928, **40**, 401–417.
16. GUILFORD, J. P., & DALLENBACH, K. M. A study of the autokinetic sensation. *Amer. J. Psychol.*, 1928, **40**, 83–91.
17. HAGGARD, E. A. Experimental studies in affective processes: I. Some effects of cognitive structure and active participation on certain autonomic reactions during and following experimentally induced stress. *J. exp. Psychol.*, 1943, **33**, 257–284.
18. HAGGARD, E. A., & ROSE, G. Some effects of mental set and active participation in the conditioning of the autokinetic phenomenon. *J. exp. Psychol.*, 1944, **34**, 45–59.
19. HUDGINS, C. V. Conditioning and voluntary control of the pupillary light reflex. *J. gen. Psychol.*, 1933, **8**, 3–51.
20. JASPER, H., & SHAGASS, C. Conscious time judgments related to conditioned time intervals and voluntary control of the alpha rhythm. *J. exp. Psychol.*, 1941, **28**, 503–508.
21. LASHLEY, K. S. Patterns of cerebral integration indicated by the scotomas of migraine. *Arch. Neurol. Psychiat.*, 1941, **46**, 331–339.
22. LEVINE, R., CHEIN, I., & MURPHY, G. The relation of intensity of a need to the amount of perceptual distortion: A preliminary report. *J. Psychol.*, 1942, **13**, 283–293.
23. MENZIES, R. Further studies of conditioned vasomotor responses in human subjects. *J. exp. Psychol.*, 1941, **29**, 457–482.
24. MILLER, J. G. Discrimination without awareness. *Amer. J. Psychol.*, 1939, **39**, 562–578.
25. MILLER, J. G. The role of motivation in learning without awareness. *Amer. J. Psychol.*, 1940, **53**, 220–239.
26. PRONKO, N. H., & KELLOGG, W. N. Reflex mechanisms and the ease of conditioning. *Amer. J. Psychol.*, 1942, **55**, 371–384.
27. PROSHANSKY, H., & MURPHY, G. The effects of reward and punishment on perception. *J. Psychol.*, 1942, **13**, 295–305.
28. RAPAPORT, D. *Emotions and memory.* Baltimore: Williams & Wilkins, 1942.
29. SANFORD, R. N. The effects of abstinence from food upon imaginal processes: A further experiment. *J. Psychol.*, 1936, **3**, 145–159.
30. SCHAFER, R., & MURPHY, G. The role of autism in a visual figure-ground relationship. *J. exp. Psychol.*, 1943, **32**, 335–343.
31. SHAGASS, C. Conditioning the human occipital alpha rhythm to a voluntary stimulus. A quantitative study. *J. exp. Psychol.*, 1942, **31**, 367–379.
32. SHAGASS, C., & JOHNSON, E. P. The course of acquisition of a conditioned response of the occipital alpha rhythm. *J. exp. Psychol.*, 1943, **33**, 201–209.
33. SHERIF, M. *The psychology of social norms.* New York: Harper & Bros., 1936.
34. SKOLNICK, A. The role of eye movements in the autokinetic phenomenon. *J. exp. Psychol.*, 1940, **26**, 373–393.

FURTHER STUDY OF THE RETENTION OF VERBAL AND MOTOR SKILLS

BY FRANCES VAN DUSEN AND HAROLD SCHLOSBERG

Brown University

Introduction

In 1944, Leavitt and Schlosberg (2) reported that retention for the pursuit rotor is clearly superior to retention of nonsense syllables. The superiority of the motor skill appeared at one day after the original learning, and was still present after 10 weeks of no practice. The differences were not dependent on the method of scoring, for they appeared in percent savings, in raw scores on the first retention trial, and even when most of the reminiscence was parcelled out. After considering the long and controversial history of the topic, Leavitt and Schlosberg felt it unwise to generalize their clear-cut results, and to conclude that motor tasks are retained better than verbal tasks. Instead, they looked for specific factors in the two tasks that would account for the differences.

Their first suggestion is the one which concerns us in the present paper. They argued that nonsense syllables, learned by the anticipation method, constitute a less integrated task than does the pursuit rotor. Hence, one would expect the pursuit rotor to be superior in resistance to disintegration during a period of no practice. The obvious way to settle the question is to use two tasks that are organized in a similar fashion. Leavitt and Schlosberg suggested the use of a verbal task such as paragraph learning which might be as well integrated as the pursuit rotor. Further thought convinced us that an attempt to match the pursuit rotor with a verbal task was easier to suggest than to carry out, especially since the pursuit rotor task is resistant to analysis into its elements. Hence, we decided to make the comparison between verbal and motor tasks which had a simple and clear-cut organization.

The present experiment was based on the method of paired associates, which can be used equally well on perceptual-motor and on verbal materials. The apparatus consisted of a large panel containing two sets of eight toggle switches, each of which was labelled with a nonsense syllable. One set of labelled switches made up the first member of each of eight pairs; the other set furnished the paired associate. Thus, the *S* had to learn a number of perceptual-motor

units, each consisting of snapping one switch, a gross arm movement, and the snapping of the second switch of the pair. The *S* simultaneously learned the syllables which appeared as labels on the paired switches. After the *S* had reached a criterion on the combined tasks, he was tested on them separately. Various *S*s were retested after 1, 7, and 28 days. It will be shown in the section on Results that there was no significant difference in the retention of the two habits.

Apparatus and Method

The apparatus was built on a 40-in.-square Masonite panel which was supported in an upright position so that the bottom edge was 36 in. from the floor. Sixteen toggle switches were fastened in two concentric circles, of eight switches each, to the front surface of the panel (see Fig. 1). The radius of the inner circle was two in. while that of the outer circle was 17 in. Two

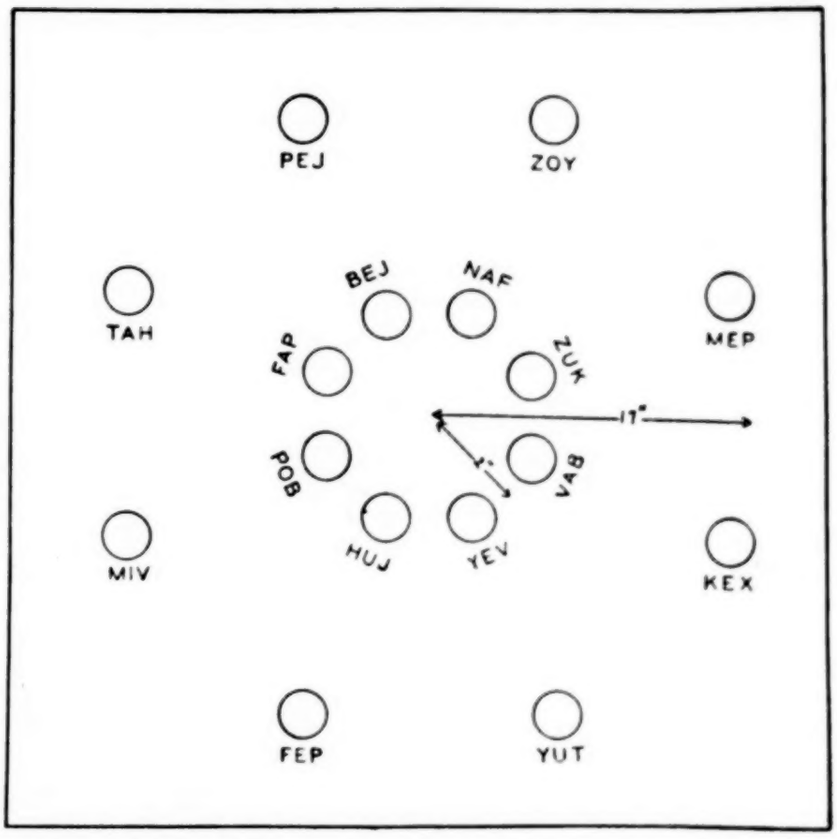

Fig. 1. A schematic diagram of the apparatus. The dimensions are not drawn to scale, but all angular relations have been maintained. The circles represent the location of the switches.

switches were placed on the circumference of the outer circle at the top of the panel, two at the bottom, two at the right side, and two at the left side. The arrangement avoided having switches located directly opposite the corners of the panel and directly opposite the midpoints of the sides of the panel. This made it difficult for the *S*s to use distinctive verbal schemes, as 'north,' or 'top,' and forced them to rely more upon a visual-motor type of memory.

Each of the switches on the inner circle was connected to an outer switch. There were, thus, eight different pairs of switches. The second switch in each pair was assigned so that the movements involved in successively turning on the inner and outer switches were as diverse as possible. The arrangement of the wiring was such that closing an inner switch, followed by closing the correct outer switch, sounded a buzzer.

A removable panel with holes through which the handles of the toggle switches projected, hooked over the front of the first panel. Metal card holders ½ in. high and two in. long were fastened to the front of the outer panel below the circular holes. Thus, there was a card holder below the handle of each toggle switch. The front panel with the card holders could be removed, leaving only the first panel to which the unlabelled toggle switches were attached. This arrangemade it possible to test the retention of the switch habit separately.

Sixteen nonsense syllables printed on cards were placed in the holders. These syllables were selected from Hull's list (1), and were, as far as possible, those with lowest association values. The pairs of syllables corresponding to the pairs of switches were: NAF-KEX; ZUK-PEJ; VAB-FEP; YEV-TAH; HUJ-MEP; POB-YUT; FAP-ZOY; BEJ-MIV. The first syllable in each pair was placed beneath the switch in the inner circle, and the other syllable in the pair was placed beneath the corresponding switch in the outer circle.

Fifty-one *S*s took part in the experiment. They were all male college students enrolled in an elementary or social psychology course. The *S*s stood squarely in front of the board at a distance from which they could comfortably reach all of the switches by gross arm movements.

On the learning trials the *S*'s task was to discover by a trial and error process the correct responses and to learn them. Turning on the eight pairs of switches in a clockwise order constituted a trial. On the first trial the *S* turned on the first of the inner switches and then one of the outer ones. If the outer switch chosen was correct, the buzzer sounded; if it was incorrect, no sound was heard. In either case, the *S* first turned off the outer switch and then the inner one. He proceeded to turn on the next inner switch and to try one outer switch to see if it would sound the buzzer for that next pair. Thus, on any one trial, the *S* had only one opportunity to find the correct outer switch for a particular inner switch. In this manner, the *S* continued working around and around the circle until all of the pairs of switches were located and learned.

The directions read to all of the *S*s informed them that their task was to learn the correct pairs of switches and the pairs of syllables corresponding to them. They were asked to pronounce aloud the syllable accompanying each switch as they turned it on. They were told that they would receive separate tests on each material at the completion of the learning.

Each *S* was given as many trials as was necessary to enable him to turn on all pairs of switches correctly three successive times. The learning took place in one session under conditions of massed practice. No intertrial rest interval was given.

After each *S* had reached the criterion of learning, he was tested on the verbal and perceptual-motor material separately. This pre-retention test was necessary to obtain a measure of the degree of learning of the materials which had been acquired simultaneously. The motor test was given on the panel which held only the toggle switches. In the absence of the syllables, the *S*'s task was to turn on correctly each pair of switches. The score was the number of pairs of switches correctly turned on. The test of the syllables was given in a different spatial situation. The syllables corresponding to the inner circle of switches were printed in a column on a card. Those corresponding to the outer circle were printed in a second column, but their order was not the original one. Starting with a syllable that had appeared on the inner circle, the *S* looked down the list of syllables that had appeared on the outer circle until he found the one which he believed had correctly completed the pair. The score was the number of pairs of syllables correctly recognized. The time elapsing after the last criterial trial and the first test averaged one minute. As soon as the *S* finished the test of one material he was immediately given the test of the other. Half of the *S*s were tested on the switches first, and half on the syllables first.

To compare the retention of the materials the total group of 51 *S*s was divided at random into three subgroups. Twenty-three *S*s were tested for retention of the materials after a 28-day interval; 16 were tested after a seven-day interval; and 12 were tested after a one-day interval. An effort was made to keep the *S*s from knowing that they were to be retested on the same material, until they actually started the retention tests. For each *S* the order of testing for retention was the reverse of that used in his pre-retention test. The method of administering pre- and post-retention tests was identical, except that a new card with a different random order of response syllables was used on each. In all cases the retention tests were given within one hour of the time of day at which the original learning had been tested.

Results

The number of trials required by the *S*s to reach the criterion of three successive errorless trials ranged from eight to 87 with a mean

TABLE I

LEARNING AND RETENTION OF SYLLABLES

The Mean and Median Number of *Syllables* Correct on the Immediate Test and on the Retention Test for each of the Three Experimental Groups

Group	Number of *Ss*	Immediate Test	1-Day Test	7-Day Test	28-Day Test
1 day	12	Mdn. 5 Mean 4.75	Mdn. 6 Mean 5.42		
7 day	16	Mdn. 4.5 Mean 5.06		Mdn. 2 Mean 3	
28 day	23	Mdn. 4 Mean 4.22			Mdn. 1 Mean 1.39

of 30.67. The verbal and motor habits had received the same number of reinforcements, so that differences among Ss in the number of learning trials should not have produced a differential effect upon the retention of the two types of material. Thus, it was not necessary to eliminate differences among the three subgroups in the mean number of trials required for learning; this was one reason for using simultaneous learning of the two types of material.

The median and mean number of syllables correct were computed for each of the three subgroups on the pre-retention test and on the post-retention test. The same measures were computed for the number of switches correct on those tests. These data are presented in Tables I and II. The means are also presented graphically in Fig. 2. The mean and median give essentially the same account of learning and retention. Since there are more statistical techniques available for dealing with the mean, this measure was used in all further treatment of the data.

The first step in the analysis of the results was to compare the three subgroups in terms of degree of original learning of each type of

TABLE II

LEARNING AND RETENTION OF SWITCHES

The Mean and Median Number of *Switches* Correct on the Immediate Test and on the Retention Test for each of the Three Experimental Groups

Group	Number of *Ss*	Immediate Test	1-Day Test	7-Day Test	28-Day Test
1 day	12	Mdn. 6.5 Mean 6.58	Mdn. 6 Mean 5.67		
7 day	16	Mdn. 7 Mean 6.56		Mdn. 4 Mean 4	
28 day	23	Mdn. 7 Mean 6.22			Mdn. 2 Mean 2.48

material. This information could not be obtained from the original learning data, since the *Ss* had the opportunity to respond to each pair on the basis of switch location, nonsense syllable, or both. Hence, it was necessary to insert the separate pre-retention tests of each material approximately one min. after the learning criterion was reached. The score made by each *S* on each type of material on this test represents the starting point of the curve of forgetting.

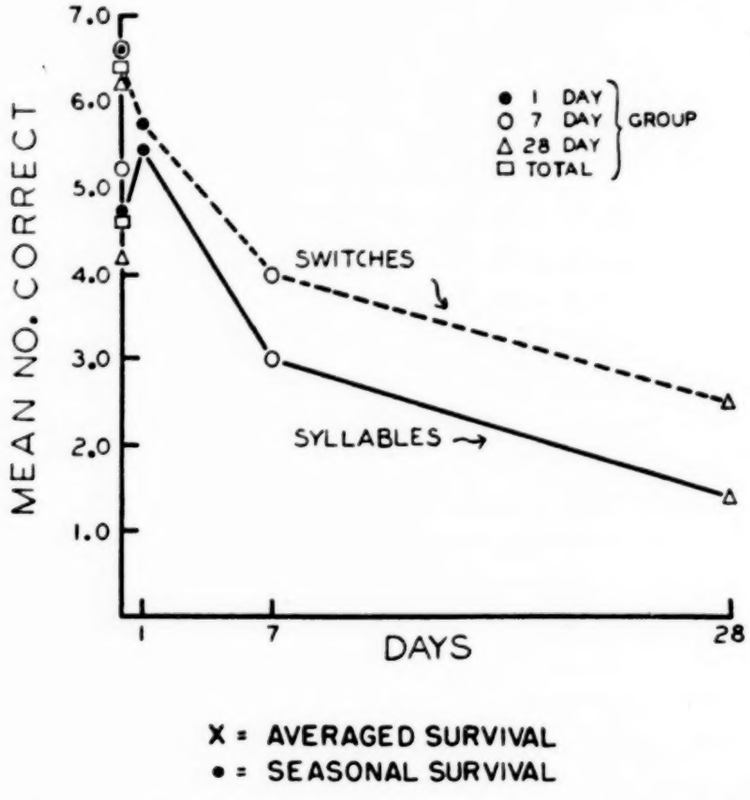

FIG. 2. The mean number of syllables and of switches correct on the retention test for the groups allotted to each of the three retention intervals. The means for the immediate test have been plotted for the subgroups separately and when combined into a total group of 51 *Ss*.

Considering the syllables first, the variance of the number of syllables correctly recognized by the various retention groups on the pre-retention test was analyzed. The *F*-ratio of the variance among the subgroups to the average variance within the subgroups was found to be not significant. Hence, the hypothesis is tenable that the three subgroups are random samples from the same population in terms of their degree of learning of the syllables.

The variance of the number of switches correctly recognized on the pre-retention test was similarly analyzed. Again the *F*-ratio of the variance among the subgroups to the average variance within the subgroups was found to be not significant. Thus, the three subgroups are comparable in terms of their performance on the pre-

retention tests of the syllables and of the switches. This fact made it possible to investigate the retention of the syllables and switches at three different retention intervals, since the three subgroups may be considered to be equated in terms of degree of original learning of each of the materials.

Next, a comparison was made between the mean number of syllables and of switches correct on the pre-retention tests. The scores of the three subgroups on each material were combined, making two distributions of 51 scores each. The mean number of syllables correct was 4.61 and the mean number of switches correct was 6.41. Since each *S* acted as his own control in learning the two materials, the *t*-test in the form for related measures was used to test the difference between these means. This test indicated that the difference in favor of the switches was significant at the one percent level of confidence. The same comparison was made for each of the subgroups separately and showed the differences in favor of the switches to be significant at either the one or five percent levels of confidence. Thus, there is a slightly greater degree of original learning of the switches than of the syllables as measued by the pre-retention tests.

The results of each subgroup on each material were investigated to determine whether or not there were significant losses during the elapsed retention intervals. The performance of a group on its pre-retention test was compared with its performance on the test after the retention interval. For the syllables, the groups with 28- and seven-day intervals showed losses which were significant at the one percent level of confidence. The one-day group showed a gain (reminiscence?) in the mean number of syllable correct after one day, but this gain was not significant at even the 20 percent level. For the switches, the 28- and seven-day groups showed losses significant at the one percent level, while the one-day group showed a loss significant at the five percent level of confidence.

The most important comparison to be made is that of the retention of the syllables and switches. On the retention tests at 1, 7, and 28 days, each group showed a higher mean number of switches correct than of syllables. Various procedures were tried to determine whether these differences in favor of the switches were significant. However, none of the procedures showed the differences to be significant. The *t*-test for related measures was chosen for presentation here, since it seems to be the best method for testing the significance of the differences in the mean number of syllables and switches correct on the retention tests. When this measure was used, none of the differences in favor of the switches was significant at even the five percent level of confidence. In Fig. 2 the retention curves for the two types of material appear to be roughly parallel, except for

the upturn of the one-day group on syllables. There is certainly no indication that the motor habit is retained better than is the verbal one.

As a matter of fact, there is even a hint that the syllable pairs are retained better than the switch pairs. The switches started out significantly superior to the syllables on the pre-retention test, but were no longer significantly superior after seven or 28 days. From Fig. 2 it can be seen that the difference on the pre-retention test is almost twice as large as that found on the seven- or 28-day tests. But these differences are absolute, and one would expect them to decrease as the level of the retention curves fall. When the two retention scores for each *S* are expressed as percentages of his pre-retention scores, the two retention curves (not shown here) are almost identical, except for the reminiscence-like hump on the one-day syllable test.

Discussion

The results of the present experiment show that paired nonsense syllables are retained about as well as are similarly organized visual-motor habits. These results are to be contrasted with those of Leavitt and Schlosberg (**2**), who found markedly better retention for the pursuit rotor, when compared with nonsense syllables.

In explaining the differences between the two sets of results, one is tempted to say that learning to actuate paired switches is not a very motor task. If this is true, the present experiment was concerned with comparing two rather verbal tasks. A similar criticism may be directed against McGeoch (**3**) and McGeoch and Melton (**4**) since the stylus mazes they used as motor tasks undoubtedly involved a strong verbal component. But actually it is quite improbable that verbal elements played a very important part in learning to actuate the switches. As indicated in the description of the apparatus, every effort was made to locate the switches so that it would be difficult to verbalize their location. They did not match the normal dimensions of up-down or right-left. They did not correspond to the cardinal points of the compass, or to the hours on a clock face. Furthermore, *S* was constantly naming the accompanying nonsense syllables as he actuated the switches. As a final check, several *S*s were asked how they learned the locations of the switches, and all denied the use of verbal designations. Hence, there is every reason to believe that the present motor task was considerably nearer the motor end of the verbal-motor dimension than was maze learning. On the other hand, it is unlikely that learning switches is as independent of verbal elements as is learning of materials of the pursuit rotor type (**5**). In other words, if nonsense syllables represent one end of the continuum,

and the pursuit rotor represents the other end, the maze would probably be fairly close to the nonsense syllables, while learning to actuate switches would be near the pursuit rotor. If the hypothesis were true, that motor materials are retained better, we should expect to find the pursuit rotor retained best, actuating switches almost as good, the maze considerably worse, and nonsense syllables the worst of the four types. But the only one of these materials that shows significantly better retention than the others is the pursuit rotor. The hypothesis seems to be false.

This conclusion is consistent with the basic assumptions underlying modern learning theory. Indeed, so strong was the supposition that motor materials are not retained better than verbal ones, that Leavitt and Schlosberg sought other explanations for their observed superiority of the pursuit rotor. The present experiment strongly supports one of their suggestions; namely, that the superior retention of the pursuit rotor depends on the organization of the task. It is extremely difficult to analyze a complicated skill like the pursuit rotor into its elements, but even superficial consideration shows that it is highly patterned and integrated. The performance of a skilled operator is a smooth and unitary performance, with many partial clues and movements blended into a whole. It is not surprizing that such a skill resists disintegration during the retention period. Its organization should not be compared with that of nonsense syllables, but with a paragraph or a poem. It has long been known that such complicated verbal units are also resistant to forgetting. The experimental comparison of the retention of pursuit rotors and poems has yet to be made, and it would not be easy to do it. Even if one could solve the problem of equivalent scoring units, there would still be the question as to whether or not the integrating factors (meanings, previous associations) that hold together a paragraph are strictly comparable to those that integrate the pursuit rotor performance.

As indicated in the Introduction, the present experiment approached the problem from the other end of the scale of complexity. The organization of the perceptual-motor task was made as simple as that of the nonsense syllables, by setting them both up according to the method of paired associates. The failure to find clear superiority in retention of the perceptual-motor task in the present experiment strongly suggests that the superior retention of certain motor skills is due to their organization, rather than to the fact that they are motor.

Summary and Conclusions

1. Fifty-one Ss learned to locate and actuate 10 pairs of switches, so connected that a buzzer sounded when each correct pair was

turned on. At the same time they learned 10 pairs of nonsense syllables, attached as labels to the switches.

2. After three correct trials, they were given a pre-retention test on the two types of materials separately, and divided into three groups. The groups came back for post-retention tests after 1, 7 and 28 days, respectively.

3. There was no significant difference in retention between the two types of materials after any of the retention intervals.

4. The actuation of the paired switches is believed to be a perceptual-motor activity, fairly free from verbal elements. Therefore, the results disprove the hypothesis that motor materials are retained better than are verbal ones if both are organized in the same manner.

5. The marked resistance of certain skills, as the pursuit rotor, is apparently largely due to their integrated nature.

(Manuscript received August 27, 1947)

References

1. Hull, C. L. The meaningfulness of 320 selected nonsense syllables. *Amer. J. Psychol.*, 1933, 45, 730–734.
2. Leavitt, H. J., & Schlosberg, H. The retention of verbal and of motor skills. *J. exp. Psychol.*, 1944, 34, 404–417.
3. McGeoch, J. A. The comparative retention values of a maze habit, of nonsense syllables, and of rational learning. *J. exp. Psychol.*, 1932, 15, 662–680.
4. McGeoch, J. A., & Melton, A. W. The comparative retention values of maze habits and of nonsense syllables. *J. exp. Psychol.*, 1929, 12, 392–414.
5. Renshaw, S., & Postle, D. K. Pursuit learning under three types of instruction. *J. gen. Psychol.*, 1928, 1, 360–367.

RATE OF RECALL AS A MEASURE OF LEARNING: I. THE EFFECTS OF RETROACTIVE INHIBITION

BY LEO POSTMAN

Harvard University

JAMES P. EGAN

University of Wisconsin

AND JEAN DAVIS

There is no one measure of learning. Changes in performance as a function of practice can be described and quantified in many different ways. We can infer that an organism has learned by examining the *end products* of his activity after varying amounts of practice or we can focus our attention on changes in the *nature of the performance* through which these end products are achieved. In the quantitative study of human learning, the main emphasis has been on end products, such as number of items correctly reproduced, recognized, reconstructed, or saved in relearning. Little attention has been paid to other behavioral changes resulting from continued practice. Students of animal behavior, on the other hand, have used number of correct responses as only one of several coordinate measures of habit strength. The *temporal* characteristics of responses, such as latency and rate of response, have been found sensitive indices of the degree of learning.

Such indices may prove equally useful in human learning. It was recently demonstrated that response latency may be a more discriminating measure of retroactive inhibition than number of items retained (**10**). This series of papers will show that *rate of response* is an equally sensitive measure of verbal learning in human *Ss*. Like latency, rate of response differentiates significantly among experimental conditions when measures of sheer amount learned or remembered fail to do so.

Rate of response is a measure that has well proved its worth in the investigation of operant conditioning. It is in terms of measures of rate that Skinner's system of behavioral analysis has been translated into quantitative terms. It is in terms of the rate at which operant responses are elicited that the strength of the 'reflex reserve' is measured (**11**). Why apply this particular measure to human learning and retention? Use of common measures in animal and human experimentation is a first necessary step toward theoretical

integration. Rate of response may well prove to be a basic index in the quantification of learning, for rate describes the process of responding as it progresses in time and is sensitive to temporary fluctuations and inhibitions.

In this first experiment we have measured the effects of retroactive inhibition on rate of recall. Retroactive inhibition is, by definition, decrement in performance due to an activity interpolated between learning and recall. The hypothesis to be tested was that an interpolated activity not only decreases the *number* of items recalled but also slows down the rate at which correct responses are given.

Procedure

There were two experimental groups and a control group. The conventional paradigm for the study of retroactive inhibition was used.

Experimental Groups: Original learning—Test—Interpolated learning—Test
Control Group: Original learning—Test—Rest—Test

The interpolated learning took place five min. after the end of the first test, and the second test was given five min. after the end of the interpolated activity.

For all groups, the original learning material consisted of 18 nonsense syllables and four three-digit numbers. All nonsense syllables had an association value of 20 percent according to Glaze's norms (4). This list was presented to *S* in random order. The randomization of the items was not complete, however. *E* followed the rule that there should never be fewer than two syllables between successive digits. The entire list of 22 items was presented to *S* nine times, each time in a different serial order. The items were printed on cards and were presented for a period of two sec. each, with three sec. between successive items. There was an interval of 15 sec. between successive series. Time was kept by means of a metronome. After *S* had seen the list nine times, he was tested by the method of free recall. *S* was instructed to recall as many of the syllables and numbers as he could. With the aid of a prearranged score sheet, *E* kept a record of the time at which each item was reproduced, dividing the recall period into successive three-sec. periods. *S* was allowed to continue until five min. had elapsed during which he could not recall anything.

The interpolated series for the first experimental group consisted of another list of 18 nonsense syllables and four three-digit numbers matching the first list. This condition will be referred to as the *syllable-digit interpolation* (S-D) since the interpolated list contained more syllables than digits. The learning procedure was identical with that used during the original learning. There was, however, no test for the interpolated material.

The interpolated material for the second experimental group comprised 18 three-digit numbers and four nonsense syllables. This condition will be designated as *digit-syllable interpolation* (D-S). In all other respects, the learning procedure was again identical with that used in the original learning.

The control group rested for a period of time equal to the duration of the interpolated learning. During this rest period, the *S*s were occupied reading selections of humorous verse.

Two types of stimulus materials (syllables and digit combinations) and two types of interpolation were used in order: (1) to gauge the differential effect of an interpolated activity on *crowded* material (18 items) as contrasted with *isolated* material (four items); and (2) to compare the retroactive effects of interpolated materials homogeneous with the original learning (syllables after syllables and digits after digits) with the effect of dissimilar interpolated material (syllables after digits and digits after syllables). It was interesting to see whether rate of recall indeed would be an index sensitive enough to bring out the differences between these various conditions of retroactive inhibition.

Each of the three groups consisted of 12 undergraduates. Half of the subjects were men students at Massachusetts Institute of Technology, the other half women students at Radcliffe College. All *S*s were volunteers and assigned to the experimental groups at random.

Results

Retention loss.—The results of the experiment will first be treated in the conventional fashion, i.e., in terms of the sheer *amount* lost due to an interpolated activity. Tables I and II list the average number of syllables and digit combinations recalled on the test and retest by the three groups, as well as the statistical significance of the mean differences between the two tests.

TABLE I

Mean Number of Syllables Recalled by the Three Groups under Different Conditions of Interpolation

The Table also Lists Differences Between Test and Retest and Significance of the Differences.

Condition of Interpolation	Mean Number Recalled			*t*	*P*
	Test	Retest	Diff.		
Syllables-digits	14.75	13.42	1.33	1.96	.05–.10
Digits-syllables	14.67	14.59	.08	.26	.80
Rest	14.00	13.84	.16	.41	.70

TABLE II

Mean Number of Digit Combinations Recalled by the Three Groups under Different Conditions of Interpolation

Condition of Interpolation	Mean Number Recalled			*t*	*P*
	Test	Retest	Diff.		
Syllables-digits	3.75	3.00	.75	2.14	.05–.10
Digits-syllables	3.92	3.42	.50	2.50	.01–.05
Rest	3.75	3.75	0	0	1.00

As Table I shows, when a list of 18 syllables and four digit combinations is followed by a list of the same composition, a significant loss in syllables occurs. There is no significant loss after a rest nor is there a significant loss in syllables when a list of 18 digit combinations and four syllables is interpolated. Turning to Table II, we find that both conditions of interpolation result in a significant loss in the number of digit combinations recalled. On the other hand, there is no loss whatsoever for the control group.

Under both conditions of interpolation significant decrements occur. But does one condition of interpolation cause a significantly greater decrement than the other? In Table III, the decrements in number of syllables under different conditions of interpolation are compared. There is a greater decrement in syllables with syllable-digit interpolation than with digit-syllable interpolation or after rest.

TABLE III

COMPARISON OF DECREMENTS IN NUMBER OF SYLLABLES RECALLED UNDER DIFFERENT CONDITIONS OF INTERPOLATION

Conditions Compared	Mean Difference Between Decrements	*t*	*P*
Syllables-digits and digits-syllables	1.25	1.67	.10
Syllables-digits and rest	1.50	1.50	.10–.20
Digits-syllables and rest	− .08*	.16	.90

* A negative entry indicates a higher value on the retest than on the first test.

The differences satisfy reasonable criteria of reliability although they are not strikingly significant. It will be recalled that there were nine presentations of both the original list and the interpolated list, which led to almost perfect learning of both lists by the majority of the Ss. When either the original learning or the interpolated learning or both are very strong, the retention loss tends to be small (**6**). Hence, the differences between decrements under the various conditions of interpolation cannot be expected to be striking. We did not intend these differences in sheer amount retained to be great, since we wanted to test the sensitivity of *rate of response* as a measure with gross differences in memory loss absent.

Comparison of the decrements in number of digit combinations recalled shows a picture similar to that obtained with syllables (Table IV). Both types of interpolation lead to a significant decrement in the number of digit combinations recalled but one type of interpolation does not seem to be more effective than the other. Both types of interpolation produce more interference than the rest condition.

Rate of recall.—We now turn to an evaluation of *rate of recall* as a measure of retroactive inhibition. Does this measure differentiate among the experimental groups even after retention losses have been taken into account, i.e., when only correctly recalled items are considered? If it does, it must be a measure of memory more sensitive than sheer number of items recalled.

TABLE IV

COMPARISON OF DECREMENTS IN NUMBER OF DIGIT COMBINATIONS RECALLED UNDER DIFFERENT CONDITIONS OF INTERPOLATION

Conditions Compared	Mean Difference Between Decrements	*t*	*P*
Syllables-digits and digits-syllables	.25	.63	.50
Syllables-digits and rest	.75	1.88	.10
Digits-syllables and rest	.50	1.79	.10

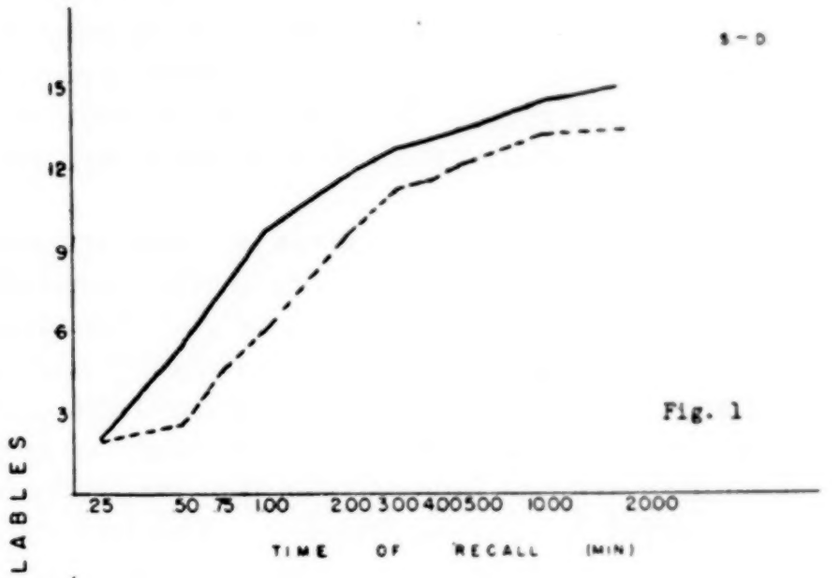

Fig. 1

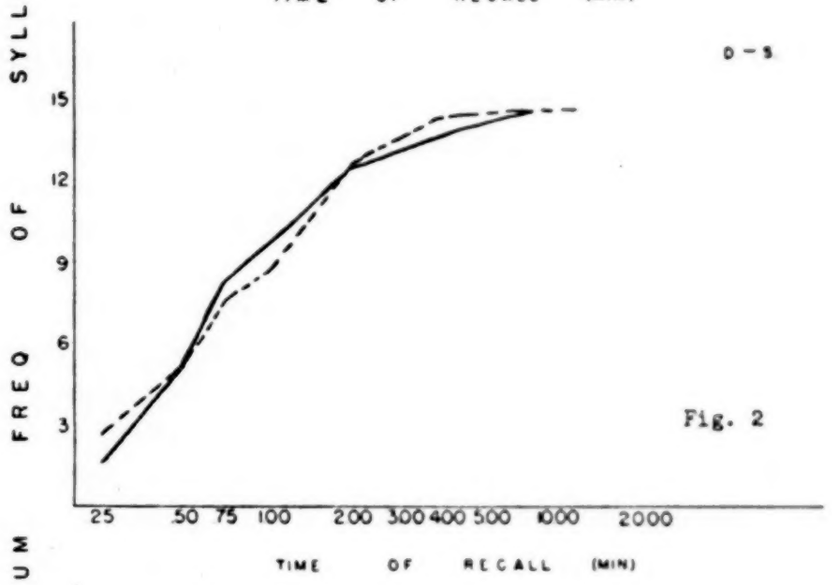

Fig. 2

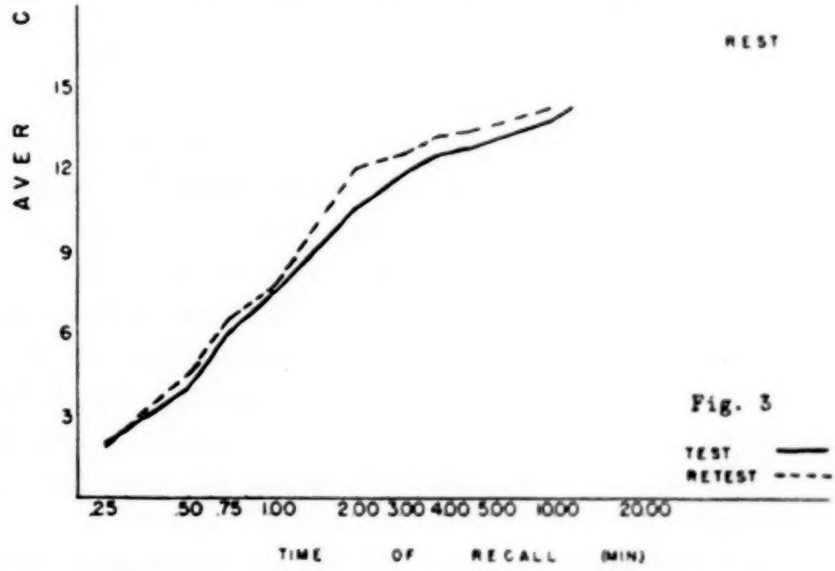

Fig. 3

FIGS. 1–3. Average cumulative frequency of syllables correctly remembered plotted against time of recall on semi-logarithmic paper

Fig. 1: Syllable-digit interpolation; Fig. 2: digit-syllable interpolation; Fig. 3: rest condition

In Figs. 1, 2, and 3, the average cumulative frequency of syllables correctly remembered on the original test and retest is plotted against the logarithm of time of recall. The cumulative frequencies represent the recorded observations; the *rate* of recall is indicated by the slope of the cumulative function.

Inspection of the three sets of curves at once reveals striking differences. With syllable-digit interpolation, the cumulative curve for retest lies consistently below the curve for the original test. There is a considerable difference in rate for the early part of the recall period. Later, the course of the two curves tends to be fairly parallel, but they never cross even though the difference in final level is small. With digit-syllable interpolation the relation between the test and retest curves is entirely different. The two curves cross twice, at .50 min. and again at about 2.00 min. The initial rate is greater on the original test whereas the rate during the later part of the recall period is higher on the retest. Finally, the retest curve of the control (rest) group lies consistently *above* the original test curve. On the retest, responses are given at a faster rate over a considerable section of the recall period, probably due to a dissipation of intraserial inhibition (**5**).

Figs. 4, 5, and 6 present comparable curves for the cumulative frequency of digit combinations recalled on the original test and the retest under the various conditions of experimentation. Here the discriminatory power of the *rate* measure is revealed even more strikingly. With digit-syllable interpolation, the test and retest curves show radically divergent trends. The slope of the original test curve is considerably steeper and the two curves never cross. Since the difference in final level (total amount retained) is small, correct responses are continued on the retest for a considerably longer time than on the original test. Especially noteworthy is the extremely slow initial rate of response on the retest, with a 'compensatory' speed-up in the later portion of the recall period. With syllable-digit interpolation, the picture is very similar, although the two curves do not diverge quite as much. The total recall periods are of approximately equal duration but the retest curve lies consistently below the original test curve. Again the slow initial rate of response during retest is remarkable. Now consider the rest condition. The original test curve and the retest curve are so similar as to be practically indistinguishable. The rates of recall on the two tests are virtually identical.

Do these divergences in rates of recall represent statistically reliable trends? To answer this question, an analysis of variance of

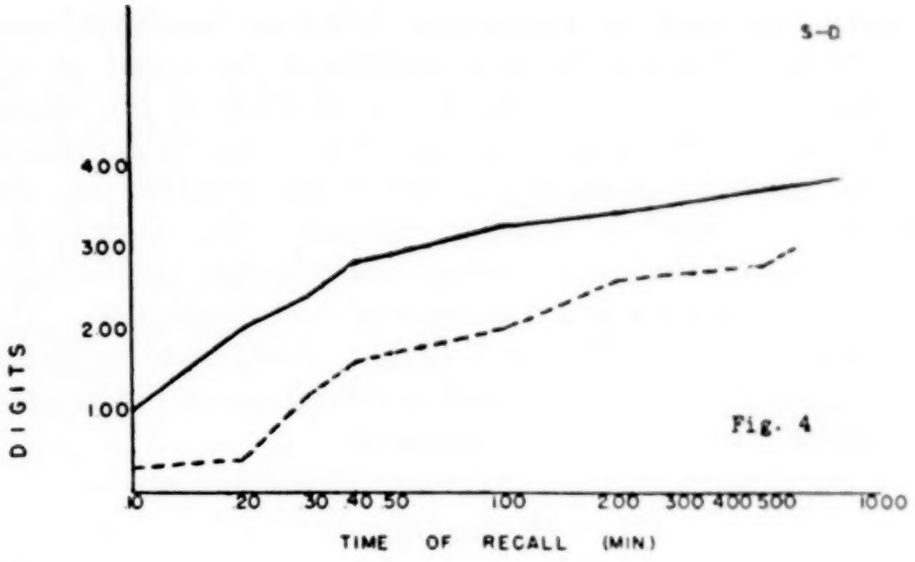

Fig. 4

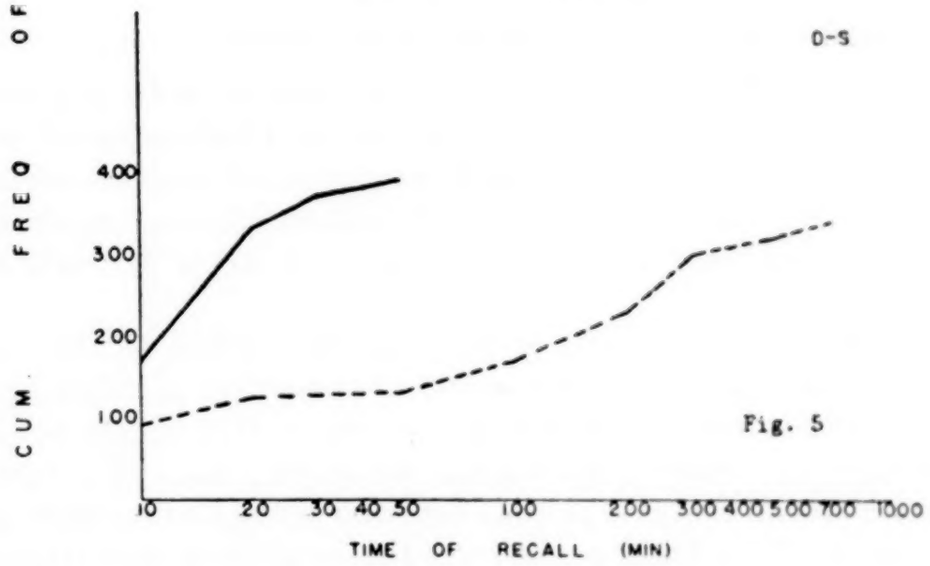

Fig. 5

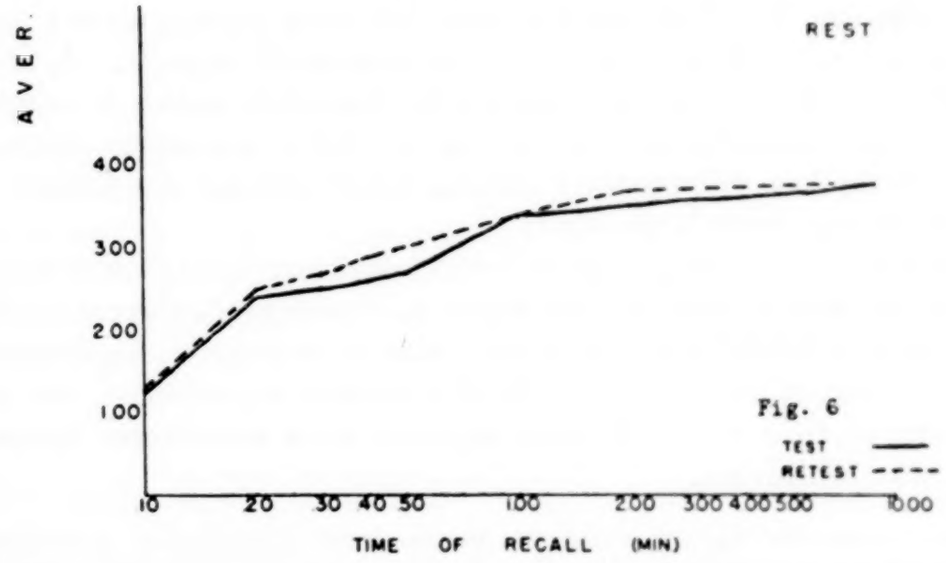

Fig. 6

FIGS. 4-6. Average cumulative frequency of digit combinations correctly remembered plotted against time of recall on semi-logarithmic paper

Fig. 1: Syllable-digit interpolation; Fig. 2: digit-syllable interpolation; Fig. 3: rest condition

retest scores, grouped in successive half-min. intervals,[1] was performed. Separate analyses were performed for recall of syllables and of digits. The two criteria of classification of the scores were condition of interpolation and time interval. For both syllables and digits, time interval yielded highly significant *F*-ratios, i.e., different amounts were recalled in different periods. The critical question for the analysis of the recall curves was whether the *interaction of condition of interpolation and time interval* was significant. This was the case for both syllables and digits. Not only were different amounts recalled from time interval to time interval but the *temporal pattern differed from condition to condition.*

Discussion

Rate of recall as a measure of retroactive inhibition

The results of the experiment establish rate of recall as a measure of memory complementary to the conventional indices based on enumeration of correct items. Indeed, rate of recall may be considered a more sensitive measure since it differenitates among experimental conditions when only correctly remembered items are taken into account.

In a previous experiment employing the method of anticipation with paired associates, it was possible to show that reaction time for correct responses may be used to gauge retroactive effects which are not detected by more conventional measures (**10**). We now find that rate of recall serves a similar function when the method of free recall is used. This finding reinforces the conclusion that retroactive inhibition (or, for that matter, any learning phenomenon) can be fully described only in terms of *all its measurable aspects.* A description of retroactive inhibition cannot be complete unless it takes into account the temporal aspects of the learner's responses, since such temporal indices differentiate among experimental conditions after other measures have been used.

The addition of temporal indices to measures of amount not only provides us with a more sensitive and more comprehensive description of retroactive inhibition, but it may also be of help in the theoretical interpretation of the process. In the present experiment, the measurements of rate of recall lend support to a two-factor theory of retroactive inhibition.

[1] The grouping into half-min. periods was to some extent arbitrary, but so would be any other grouping of a continuous variable. Grouping into periods as long as half a minute makes the measures independent of fleeting fluctuations.

Support for a two-factor theory of retroactive inhibition

The case for a two-factor theory has been argued by Melton (**7**). The phenomenon of retroactive inhibition may be regarded as the resultant of two distinct processes: (1) unlearning of the original material or skill during the interpolation and (2) competition between the effects of the original and interpolated learning during the recall test. The two types of measures presented here—measures of amount lost and rate of recall—may be regarded as measures of these two processes. Those aspects of the original activity which have been *unlearned* cannot be evoked on the retest: unlearning leads to decrements in the *amounts* retained. Other aspects of the original learning are not unlearned, but suffer competition from the more recent interpolated activity. Such items are potentially available for evocation but their recall is 'disturbed' (**8**). Such disturbance of recall—contrasted with true forgetting by unlearning—should manifest itself in a slower rate of recall. Disturbance (competition) has to be overcome before a response occurs.

Let us consider our recall curves in the light of these theoretical considerations. The curves do not differ much in final level. With nine presentations of each list, both original and interpolated activity were overlearned, and there was little unlearning due to interpolation. On the other hand, interpolation results in a slower rate of recall. Rate of recall seems to provide us with an operational definition of that part of the retroactive process which results in competition among incompatible responses both available for recall. Rate is a measure of the disturbance of recall. A parallel suggests itself to operant conditioning where rate reflects disturbances in the proportionality of reflex strength and reflex reserve (**11**).

Some inhibitory effects of interpolated activities are only transient and are dissipated in time (**6**). Our recall curves bear out this generalization. In several of the retest curves (Figs. 2, 4, 5) we find initial depressions of the rate followed by accelerations. As the recall period continues, disturbing factors are overcome to an increasing extent. Again a parallel to operant conditioning suggests itself. There are many instances in which depression of the rate (e.g., by emotional factors such as punishment) is followed by a compensatory speed-up in response elicitation (**2**, **11**).

Interserial competition

Can measures of rate help us to gauge the exact nature of the disturbance which, at least temporarily, affects responses on the test? We have already referred to Melton's hypothesis of response competition. Examination of the recall curves obtained after inter-

polation of similar and dissimilar materials confirms the view that interserial competition is, at least in part, responsible for the disturbance of recall.

Interpolation of items similar to the original learning material slows down rate of recall considerably more than interpolation of dissimilar material. Comparison of Figs. 1 and 2 shows this effect for syllables; Figs. 4 and 5 demonstrate the same finding for recall of digits. Interpolation of a considerable amount of similar material establishes a large number of responses which, by virtue of their similarity, will compete with the original learning material. The lower the discriminability among the lists, the greater will be the competition (cf. 3). This result is, of course, to be expected according to the Skaggs-Robinson hypothesis Our measurements of rate suggest one way in which the factor of similarity may operate in retroactive inhibition, *viz.*, by causing interserial competition and slowing down the rate at which correct responses are given.

Crowding and isolation of materials

Degree of retroactive inhibition is determined not only by the nature of the interpolated activity but also by the characteristics of of the original learning task. Among the many parameters on which susceptibility to interference depends is 'crowding' or 'isolation' of the materials learned (1, 9). Ever since the experiments of von Restorff, it has been clear that an isolated item in a homogeneous list is learned far more quickly than any one of the homogeneous or 'crowded' items. However, an isolated item, standing out like a figure from a homogeneous ground, is not only learned more quickly but is also more available for interference. Our measures of rate verify this deduction. Comparison of Figs. 1 and 4 tells the story. These figures show the recall curves for syllables and digits when a list of 18 syllables and four digit combinations is followed by an interpolated list of identical composition. The syllables are crowded, the digits are isolated items. The relative number of similar interfering items is the same—18 syllables followed by 18 syllables and four digit combinations followed by four digit combinations. Yet the depression in rate appears to be considerably greater in the case of digits than in the case of syllables.[2] We conclude that there was probably more interserial competition among the 'isolated' digits than among

[2] This statement is made on the basis of visual inspection only. Statistical comparison of the decrements is inadvisable in view of the great discrepancy in the number of items. Further experimentation will be necessary to establish this point since our results are somewhat at variance with the findings of Buxton and Newman (1) who report that isolated items in a homogeneous series are retained better than crowded items over considerable periods of time. These investigators did not, however, use interpolated activities similar to the original learning task, so that their results are not strictly comparable to ours.

the 'crowded' syllables. Variations in rate of recall reflect this differential effect of crowding and isolation of stimulus materials.

Summary

Amount remembered may not always be a sufficiently sensitive measure of human memory. This paper shows that rate of recall is a sensitive index which differentiates among different degrees of retroactive inhibition over and above conventional measures of retention loss.

Two experimental groups and a control group were used in the usual design for experiments on retroactive inhibition. The experimental groups learned an interpolated task five min. after original learning while the control group rested.

The original learning series for all groups consisted of 18 nonsense syllables and four three-digit numbers. For one group, the interpolated material consisted of 18 syllables and four three-digit numbers; for the other group, 18 three-digit numbers and four syllables. The lists were presented nine times in random order. The *Ss* were tested by the method of free recall at the end of the original learning and retested five min. after the interpolated activity. The speed with which responses were given was recorded on both tests.

When *correct* responses are considered, i.e., after retention losses have been taken into account, rate of recall differentiates significantly among the experimental conditions. Retroactive inhibition manifests itself in a slow-down of the rate of recall after interpolation. Interpolation of items similar to the original learning material causes a greater depression in the rate than interpolation of dissimilar material. Measured by variations in rate of recall, 'isolated' materials (digits) are more susceptible to interference effects than 'crowded' materials (syllables).

Rate of recall not only provides an additional sensitive measure of memory but is also useful in the theoretical interpretation of retention. It is a measure especially relevant to the investigation of temporary disturbances of recall.

(Manuscript received August 27, 1947)

References

1. Buxton, C. E., & Newman, E. B. The forgetting of 'crowded' and 'isolated' materials. *J. exp. Psychol.*, 1940, **26**, 180–198.
2. Estes, W. K. An experimental study of punishment. *Psychol. Monogr.*, 1944, **57**, No. 3.
3. Gibson, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196–229.
4. Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–267.

5. HOVLAND, C. I. Experimental studies in rote learning theory. I. Reminiscence following learning by massed and by distributed practice. *J. exp. Psychol.*, 1938, **22**, 201–224.
6. MCGEOCH, J. A. *The psychology of human learning.* New York: Longmans, Green and Co., 1945.
7. MELTON, A. W., & VON LACKUM, W. J. Retroactive and proactive inhibition in retention: evidence for a two-factor theory of retroactive inhibition. *Amer. J. Psychol.*, 1941, **54**, 157–173.
8. MÜLLER, I. Zur Analyse der Retentionsstörung durch Häufung. *Psychol. Forsch.*, 1937, **22**, 180–210.
9. NEWMAN, E. B. Effect of crowding of material on curves of forgetting. *Amer. J. Psychol.*, 1939, **52**, 601–609.
10. POSTMAN, L., & KAPLAN, H. L. Reaction time as a measure of retroactive inhibition. *J. exp. Psychol.*, 1947, **37**, 136–148.
11. SKINNER, B. F. *The behavior of organisms: an experimental analysis.* New York: Appleton-Century, 1938.

THE SEQUENTIAL ORDER OF CONCEPT ATTAINMENT

BY BERNICE M. WENZEL AND CHRISTINE FLURRY

Barnard College, Columbia University [1]

Introduction

Heidbreder's latest work (**4**, **5**, **6**, **7**) has been concerned with development of a new, 'dynamic' theory of cognition which has grown out of a long series of experiments. For a full discussion of the theory reference should be made to two of Heidbreder's recent publications (**5**, **7**). The chief result of her experiments has been a consistency of the order in which certain concepts are attained. The earliest concepts to be formed are those of concrete objects, next come those of spatial forms, and finally those of number.

The experiment cited to illustrate this order of concept attainment was ostensibly one on memorizing in which the *S*s learned the nonsense-syllable names of 16 successive series of drawings. A series consisted of nine drawings and all of the 144 drawings were different. There were only nine different names, however. Three names were applied to drawings of concrete objects, three to drawings of geometrical forms, and three to drawings of certain numbers of things. It was possible for the *S*s to discover the regularity, but they were not informed. An *S* was credited with attaining a given concept when he correctly named a drawing he had not seen before. The number of the series in which this occurred was recorded as the measure of attainment for that particular concept.

Definitions of these particular concept categories rests on what can best be called a common-sense basis. It is admitted that such a procedure can become dangerous by avoiding issues. In this case, however, the boundaries seem clearly drawn. A concrete object refers to "a material body in the external world, typically visible, tangible, and manipulable. Concrete objects are taken, in the first instance, as unreflectively and as uncritically as possible—as the 'things' of common sense . . ." (**5**, p. 2). The class of spatial forms includes any shape which can be represented in at least two dimensions [2] and which is essentially a nonsense figure. It may have a

[1] The research for this paper was performed at H. Sophie Newcomb Memorial College, Tulane University.

[2] It has been hypothesized, and Heidbreder agrees, that concepts of three-dimensional forms would be attained along with concepts of concrete objects, since one of the main explanations for the order is the relative manipulability of the stimuli. Some unpublished data collected by Weitz and Flurry (**11**) bearing on this hypothesis do not wholly support it.

specific meaning for a given individual due to past experience but it is not intended to represent any meaning generally. It is adaptable to a variety of materials and is not typically visible, tangible, or manipulable except as embodied in a concrete object. By number is meant a direct expression of quantity.

There is no attempt to generalize from these three classes to all possible conceptual stimuli. Once the apparently distinct progression within these three groups is clearly established, much experimentation and theorizing will be needed to determine whether there may be additional groups, how many they are and what they include, whether they are primary or secondary, and how to classify the multitude of stimuli available to us.

Since Heidbreder's original experiments were performed, others have been done by experimenters under her direction. Three of these deal with the position of color in the order of concept attainment. Overstreet (9) used objects, colors, and forms; Bensley (2) used objects, colors, and numbers; and Ivy (8) used forms, colors, and numbers. Objects, forms, and numbers were found to bear the same relation to each other that Heidbreder originally found, but color had a variable position. Overstreet found that, in her *S*s, concepts of objects were formed before concepts of spatial forms, but the difference was not statistically significant. Color came in no particular order, being mixed with concepts of the other two types. Bensley found significant differences in the number of trials required to attain concepts of objects, colors, and numbers, in that order. Ivy found significant differences in the time required to attain concepts of forms, colors, and numbers, in that order. Twenty-five *S*s were used in each of the three experiments just described. The procedure used was practically the same as that used in Heidbreder's original experiments.

Certain disadvantages accrue to this procedure, from the point of view of the *S* as well as the *E*. In the first place, since the experiment is 'ostensibly one on memorizing,' the *S* does not know from the beginning that he is to form concepts. He must, therefore, change his approach to the situation at some time during the experiment; otherwise he goes through memorizing each series separately without forming concepts. This latter response apparently was not exhibited by the large majority of *S*s. Since introspections are not encouraged during the course of the experiment proper, the exact point at which this reorientation takes place, or begins to take place, is not known. Moreover, Reed (10) has found that concepts are learned more readily and retained longer if the *S* is told in his instructions that concepts are to be formed. Another disadvantage lies in the criterion for the measurement of 'C.A.' (concept attained), which was not the point at which the *S* could define the concept, but the point at which he began making consistently correct responses on first sight of a picture. This point may have been a guess, or merely known to be correct by process of elimination. A third disadvantage, an outcome of using quantitative analysis exclusively, is that not all *S*s attained all nine concepts correctly. It was

impossible to know when the *S* had formed an adequate concept until tests were given after the experiment.

In a preliminary experiment carried out by the present authors, the purpose was to verify Heidbreder's findings under slightly different conditions. Six concepts were used—two each of objects, forms, and numbers. The general results of this experiment indicated a significant difference between the number of trials required to form concepts of objects and concepts of forms, and between concepts of forms and of numbers. The order was, then, as predicted—first, objects; second, forms; and last, numbers. One of the forms used was that of a triangle which came out significantly later than that of the other form, a circle. Two explanations for this result are offered. The first is based on a Gestalt principle stated by Bender, that "in the developing mentality, spheres (in the case where three dimensions are used) or circles are biologically better perceived than straight lines" (1, p. 200). The second explanation is in terms of the nature of the triangle. The form involves a number of sides, *viz.*, three. Consequently a concept of a 'three-sided figure' would be formed at approximately the same point as a number, 'three' for example. In regard to the second explanation it can be said that it is known that animals which can make response to a triangle as distinct from other forms apparently are unable to count or do not appreciate numbers as such.

The Problem

The main purpose of the present experiment was to test further Heidbreder's finding of the sequential order of concept formation by a modification of the original procedure. Secondarily an attempt was made to test the 'number-of-sides' and 'goodness-of-form' explanations for the late attainment of the concept of 'triangle.'

Apparatus

Two sets of concepts were used. Each set was presented in 18 series of drawings, a series consisting of six drawings each representing a certain manifestation of one of the six concepts. Each concept was represented once in each series, three times in each of the six positions. No single drawing was used more than once in each set. The pictures were simple line drawings without shading done in black. The style was identical to that of Heidbreder's pictures, samples of which have been reproduced in the literature (6, p. 182). They were prepared by hand on two strips of off-white paper 1373 cm. long and 5.6 cm. wide. Each strip contained the stimuli for one set of concepts. The strips were equally divided into 18 units of 13 squares five cm. by five cm. Each of these units constituted a series of six drawings, the first 12 squares containing alternately a picture and the nonsense syllable naming it. The drawings and syllables were centered in the squares and took up most of the area. There was a blank space at the end of each series so that *E* and *S* might know that a new series was about to start. A narrow blank area at each end allowed for fastening the strip around the presentation apparatus.

Each set of six concepts contained two concepts of concrete objects, two of spatial forms, and two of numbers. These are identified in Table I.

The objects chosen to represent the concrete object category were all familiar objects. It cannot be doubted reasonably that houses, hats, leaves, and trees have a high degree of meaning for most observers and that they are tangibly present in reality. In the form category the circle and the pentagon were used in Set I in order to test the hypothesis that the concept of a many-sided polygon is formed along with the concept of numbers. In Set II, the arch and jagged line were used to test the hypothesis that concepts of curved figures are developed more easily

than those of angular forms. These samples of spatial forms meet the requirements for the class as described above. Use of the term 'arch' is not meant to imply that an object is being represented. It is simply the most convenient way to refer to the shape in question, an elongated semicircle. The numbers were chosen so that it would be possible to see whether the numbers close together, five and six, could be distinguished from one another as readily as the more widely separated ones, four and seven. Certainly there can be no dispute that four, five, six, and seven are direct expressions of quantity.

Since the aim of the experiment was to disprove the hypothesis that no difference exists in the ease with which the three types of concepts are evoked, the concepts used were obviously not considered equated. They were considered representative of their classes, however, on the basis of comparison with data from other experiments. Anticipating the results, it can be said that the average number of trials required to attain these concepts is comparable to figures obtained for the same classes by other investigators and by the present investigators in preliminary work.

TABLE I

CONCEPTS USED, GIVING THE IDENTIFYING NONSENSE SYLLABLES AND THE APPROPRIATE CATEGORY

Set I			Set II		
Syllable	Concept	Nature of Concept	Syllable	Concept	Nature of Concept
YEM	house	concrete objects	CEF	bird	concrete objects
BOF	hat		ZEG	leaf	
ZIF	circle	spatial forms	YIL	arch	spatial forms
KEB	pentagon		KEZ	zigzag line	
WEZ	four	numbers	TOV	five	numbers
DAX	seven		MIB	six	

Apparatus for presenting the pictures to the *S*s consisted of a vertical black screen 18 in. square with a five-cm. square opening to permit one drawing or one nonsense syllable to be exposed. Two rollers were attached to the back of the screen, one above and one below the opening. Both rollers were free to move and a handle could be attached to the lower roller. One of the two strips of drawings was wound around the upper roller and the end fastened to the lower roller. This operation caused the first square to appear in the window. The window was covered until the experiment actually began. Subsequent presentations were manually controlled by turning the handle a specified amount, causing the strip to move the extent of one square and to wind further on the lower roller. At the conclusion of a session the whole strip could be rewound onto the upper roller very simply so as to be prepared for the next session. The similarity to the method of advancing roll-film in an ordinary camera is obvious.

PROCEDURE

The experiment was conducted in individual sessions. The *S* was seated at a convenient distance (approximately two ft.) before the opening in the presentation apparatus and was given the following instructions:

"You are to be presented with a series of six pictures. After each picture will be shown a nonsense syllable which is to be associated with it. You are to repeat the syllable when it is exposed. The same six nonsense syllables will be used throughout the experiment, but different pictures will be shown. Each syllable is always associated with drawings of a given type. After you have seen the first six drawings, try to anticipate the syllable whenever you see drawings of the type to which it is applied."

Each drawing and each syllable were exposed for three secs. Timing was controlled by use of a stopwatch. In all series after the first, *S* was to attempt to anticipate. If he did so, the exposure of the syllable served as immediate information concerning the correctness of his response; if he did not, the syllable served to prompt him. Each time that *S* anticipated correctly, *E* asked, 'Why?' This enabled *E* to observe the various hypotheses which *S* tried out in the process of concept development. If *E* could not record *S*'s response during the six sec. in which the picture and the syllable following were exposed, then the syllable and not the picture was shown for one or two sec. longer. *S* was aware that *E* was recording his responses. *E* continued to present series after series until *S* had adequately defined each of the six concepts in response to *E*'s question. The number of the series in which *S* correctly *defined* a concept was taken as the measure of 'concept attained.' *E* experienced no difficulty in judging the definitions. They were expected to meet certain standards which were set up somewhat arbitrarily. WEZ ('four'), for example, was considered correct only when it was described as 'four figures,' 'four objects,' 'four things all alike,' or something similar. A response of 'two pairs' was judged incorrect. The definitions for YIL ('arch') and KEZ ('zigzag line') were accepted if the shape was adequately described, in words or by a picture. *S*'s responses, correct or incorrect, were recorded, as well as his verbalizations of the concepts and any introspections which he gave spontaneously.

The *S*s were 40 college women, divided into two groups. Each group learned one of the two sets of concepts. It was originally intended to have each *S* learn both sets of concepts, but it was found that after having learned one set, the *S* learned the next in only three or four trials. Therefore it was decided to have half the *S*s learn only Set I while the other half learned only Set II. The difference between the average number of trials required for the *S*s learning Set I to attain all six concepts and the average number for Set II was less than one trial and was not significant. Averages for each concept and each type of concept for the two groups of *S*s (Set I and Set II) can be seen in Table II.

Results and Discussion

The average number of trials required to attain each concept and each of the three types of concepts is shown in Table II. Results for Sets I and II are presented individually, consisting of the mean number of trials required by the 20 *S*s to attain each of the six individual concepts in each set as well as a mean for the three classes of

TABLE II

Mean Number of Trials Required to Attain Each Individual Concept and Each Type of Concept

Set I: 20 *S*s		Set II: 20 *S*s		Sets I and II: 40 *S*s
Concept	Mean Trials	Concept	Mean Trials	Mean Trials
Concrete objects		Concrete objects		
YEM (house)	2.75	CEF (bird)	2.75	
BOF (hat)	3.35	ZEG (leaf)	3.50	
Both	3.05	Both	3.13	3.09
Spatial forms		Spatial forms		
ZIF (circle)	6.15	YIL (arch)	6.50	
KEB (pentagon)	7.25	KEZ (zigzag line)	6.60	
Both	6.70	Both	6.55	6.63
Numbers		Numbers		
WEZ (four)	11.85	TOV (five)	13.55	
DAX (seven)	14.15	MIB (six)	13.35	
Both	13.00	Both	13.45	13.23

concepts represented twice in each set. Finally, the table presents the results of the two sets combined since the data for the two did not differ significantly. Table III contains the values of *t* for the differences between the means for each type of concept in Set I, Set II, and for the two sets combined.

The concept definitions, both incorrect and correct, and the comments of the individual *S*s constitute the qualitative results and are not reproduced here. These data will be discussed below.

TABLE III

SIGNIFICANCE OF DIFFERENCES (*t*) BETWEEN MEANS FOR DIFFERENT TYPES OF CONCEPTS

	Set I		Set II		Sets I and II	
	Forms	Numbers	Forms	Numbers	Forms	Numbers
Objects	3.88	6.77	8.55	11.35	7.22	12.52
Forms		3.48		3.03		8.47

The means indicate that the concepts were attained in a definite order—concepts of concrete objects first, spatial forms next, and numbers last. This order is statistically significant. Intra-category differences are significant in no case except for those between CEF and ZEG in Set II, and ZEG and YEM, the latter in Set I.

The obtaining of qualitative, as well as quantitative, results in this experiment provided two distinct advantages: first, it enabled *E* to observe the various hypotheses tried out during the process of concept development; second, it made possible a uniform criterion for the establishment of the point at which a concept was attained. The continuation of the experiment until each *S* had verbalized all six concepts made the quantitative results more valid for comparison, since there were no instances in which a concept was never attained.

The quantitative data show the same trend as Heidbreder's results. The mean number of trials required to attain concepts of concrete objects was 3.09, spatial forms 6.63, and numbers 13.23. The differences between our averages and the lower ones reported by Heidbreder for the last two categories are considered to be due to the more rigid criterion of 'concept attained' which was adopted in the present experiment. The differences between the various types of concepts are statistically reliable at the .01 level. There were no outstanding differences between individual concepts within a given type even when the two sets were combined. This fact means that, not only was there a consistency of sequence from one set to the other, but also there was agreement in the number of trials required to attain each type in the two sets. In other words, results of the

two sets were alike absolutely as well as relatively. The similarity can be seen readily in Table III where the mean number of trials required to attain each type of concept is given for the individual sets.

The trend revealed by the figures in Table II is characteristic of the 20 *Ss* when they are combined in a group and treated as one. It is of interest to know whether the order of attainment for each individual repeats the general trend. The raw data show that there is very little divergence among the members of each group. An individual was considered to be in conformity with the group if she attained the two object concepts first, then the two form concepts, and finally the two number concepts. If she happened to attain any two of the similar concepts in reverse sequence from that of the group it was not counted as a deviation. The general order is what is of interest. In Set I, six of the 20 *Ss* attained the concepts in a different order. There was no striking consistency among the variations, indicating that no one significant factor was operative in the exceptional cases. In Set II all of the 20 *Ss* attained the six concepts in the same order. There were no deviations from the group sequence.

Two of the 18 intra-category differences were reliable at the .05 level, *viz.*, ZEG with CEF and with YEM. All of these concepts were of concrete objects. The first two concepts occurred in Set II, the last in Set I. The separation of ZEG from the other concepts of concrete objects may possibly be explained by the fact that in the first two drawings the concept of 'leaf' was represented by leaves growing on and falling from a tree. Many *Ss* formed the concept 'tree' for ZEG immediately and later had to change it to 'leaf,' causing a delay in the appearance of the correct definition.

It was often noted that an *S* would anticipate a syllable correctly on several successive series without having attained an adequate concept. This ocurred most often in the case of numbers, after several concepts had already been attained. When asked why after a correct response, the *S* would answer, 'I don't know,' 'just a guess,' or 'by process of elimination.' It is possible that this type of response is the basis for the 'non-verbal' formation of concepts reported by other experimenters. It occurs in the absence of any hypothesis and can probably be described as spectator behavior. The *S* knows that a certain picture must be DAX because it is not any of the concepts with which he is already familiar, but he has yet to discover the property which is common to all DAX's.

All of the *Ss* gave evidence of both trial and error and insightful behavior. The trial and error process was not a mere hit or miss activity, even at first. The *Ss* attacked the problem directly, trying out various hypotheses in a more or less orderly fashion. Only one *S* abandoned an hypothesis after it had proved to be adequate.

This same *S* showed a tendency to cling to certain other hypotheses even after they had proved inadequate. In general, the *S*s tended to respond first to the more 'thing-like' aspects of the pictures. This is seen particularly in the early formation of concrete objects, although the data obtained from a few *S*s indicate that they were responding in terms of the situation as a whole rather than to the discrete 'things' making up the picture. The early confusion of ZEG and CEF, both 'having something to do with nature,' is an example of this generalized, and as yet unexplained, type of hypothesis which had to be narrowed down. The important fact is that regardless of the emphasis placed on the use of concepts by the instructions, the categories remained distinct. Even when the *S*s were told to search for similar types of drawings, they found the similarity among the pictures of concrete objects first, then among the spatial forms, and last among the numbers.

Spatial forms and arrangements were often tried out as hypotheses for the syllables representing a number before the concepts of numbers were attained. Once a concept of one number had been formed, however, the other number was not long in following. There was a great deal of overlapping in the attainment of concepts of 'five' and 'six' (Set II), but in Set I 'four' was always attained either previous to or along with the concept of 'seven.' The reports of several *S*s indicate that pairedness, a readily perceptible situation, could be used as a basis for the development of the concept 'four,' which is a pair of pairs. Several *S*s suddenly grasped the concept of 'five' when the picture of a hand—five digits—was shown.

In a general way, the order in which the three types of concepts were formed can be best explained in terms of their manipulability. Manipulability in these instances is potential only, of course, since all of the representations were two-dimensional and nonmanipulable except as strips of paper. An actual object can be picked up and handled and taken apart; a form, a circle for example, cannot be picked up as such, but can be traced with the eyes or fingers; a number cannot be handled by the motor equipment at all except by means of more concrete symbols.

The number of trials required to attain the concept KEB (pentagon) was not significantly different from the number required to attain concepts of other spatial forms, although the mean is higher than the others. The *S*s' responses indicate that there was a high association value of the syllable KEB since the first setting of a pentagon was a five-sided kite. The *S*s first associated the syllable beginning with K with the kite which also began with K, later with something 'shaped like a kite,' and still later with something shaped

like a specific kite which had five sides. This association apparently made the concept KEB less difficult to attain. The other explanation of the late development of the concept of triangle, based on Bender's statement that curves are biologically better perceived than straight lines, is not in accord with the results of this experiment. The concept KEZ in Set II was a jagged line and the concept YIL a semicircle or arch, yet there was no significant difference between the means for KEZ and YIL. The position of the many-sided polygon is still not clear. There is no evidence that goodness-of-form makes for less difficulty in the attainment of concepts. The 'circle' concept was developed much later than the concept of 'house,' yet the circle far surpasses the house in degree of goodness-of-form. The concept of the rounded arch appeared at just about the same time as that of the jagged line.

It is possible to conceive of the development of children's language following an order such as has been described here. The child learns first the names of concrete objects which he can handle and which have meaning and purpose for him. Later he learns the names of spatial forms, probably in terms of objects which he already knows. 'Round' means 'shaped like a saucer' or 'like a ball.' Still later he learns the names of numbers. These, too, are represented to him concretely, in terms of his own digits or familiar objects. Three apples and one apple equal four apples; therefore, three and one equal four.

Summary and Conclusions

An experiment was performed to test Heidbreder's finding that concepts are formed in a definite order—concrete objects first, spatial forms next, and numbers last. In addition, a test was made of two possible explanations for the delayed attainment of the concept 'triangle' in a preliminary experiment. The first reason suggested was that forms which can be broken up readily into a 'number of sides' may be perceived more as number than as form. The other was that goodness-of-form favors early attainment while a relative lack of it delays attainment. The experimental conditions were somewhat different from those investigations previously reporting this order. Six concepts were to be formed, two in each of the categories mentioned above. The *S*s were shown line drawings representing various aspects of each concept. The concept 'leaf,' for example, was represented by pictures of many different kinds of leaves and many different conditions. Each concept was named by a nonsense syllable.

The anticipation method of learning was used. A drawing was exposed, by means of an exposure device, for three sec. It was

followed by its nonsense-syllable name, also exposed for three sec. The *Ss* were instructed that they were to apply each syllable to a certain type of picture. Thus, they were set to form concepts rather than merely to learn names. There were two sets of six concepts each. A set consisted of 18 series of six drawings each, representing the individual concepts. No drawing was shown twice. All *Ss* formed the six concepts within the limit of 18 series. The criterion for attainment was correct definition of the concept in addition to correct naming or use. *E* recorded for each drawing the syllable anticipated by *S* and his reason. Twenty college women were *Ss* for each set.

Results consist of the numbers of the series in which the concepts were defined correctly. When the individual *Ss*' results were averaged for each of the concepts it was found that the predicted order was correct. All of the concepts of concrete objects were formed earlier than those of spatial forms, while the latter were formed earlier than those of numbers. All of the inter-category differences were significant at the .01 level and only two of the 18 intracategory differences reached significance at the .05 level. The following conclusions are justified:

1. The results revealed the same general tendency as other similar experiments in regard to the sequence of attainment of the three types of concepts tested. Concepts of concrete objects were attained first, spatial forms next, and numbers last, forming three clear-cut categories.

2. There were slight differences in order of attainment of individual concepts within a given type, but these differences were not statistically significant for the number of *Ss* used except in two cases when there were significant differences between concepts of concrete objects. This finding is in support of Heidbreder's statement that the order is continuous rather than one of discrete steps.

3. The position of the five-sided polygon in the progression from form to number could not be clearly established because of an associational tendency with the syllable used which facilitated early attainment of that particular concept.

4. There is no evidence that goodness-of-form facilitates early formation of a concept, since the concept 'arch' was not formed significantly earlier than that of 'jagged line.'

5. In an experiment of this type, both trial and error and insightful behavior are observable. The approach is a direct one, involving the use of various hypotheses until an adequate one is reached. More participant than spectator behavior was observed.

6. Among the college students studied, there are individual differences in number of trials required to learn each concept, and a few differences in the order in which the types of concepts are attained.

(Manuscript received September 3, 1947)

References

1. Bender, L. Gestalt principles in the sidewalk games and drawings of children. *J. genet. Psychol.*, 1932, **41**, 192–210.
2. Bensley, Mary L. The production of concepts of objects, colors and numbers. Unpublished Master's thesis, Wellesley Coll., 1945.
3. Heidbreder, Edna. A study on the evolution of concepts. *Psychol. Bull.*, 1934, **31**, 673, (Abstract)
4. Heidbreder, Edna. The attainment of concepts—a psychological interpretation. *Trans. N. Y. Acad. Sci.*, 1945, **7**, 171–188.
5. Heidbreder, Edna. Toward a dynamic psychology of cognition. *Psychol. Rev.*, 1945, **52**, 1–22.
6. Heidbreder, Edna. The attainment of concepts. I. Terminology and methodology. *J. gen. Psychol.*, 1946, **35**, 173–189.
7. Heidbreder, Edna. The attainment of concepts. II. The problem. *J. gen. Psychol.*, 1946, **35**, 191–233.
8. Ivy, Margaret L. The production of conceptions of spatial forms, colors and numbers. Unpublished Master's thesis, Wellesley Coll., 1945.
9. Overstreet, Phoebe L. The production of conceptions of objects, colors and forms. Unpublished Master's thesis, Wellesley Coll., 1944.
10. Reed, H. B. Factors influencing the learning and retention of concepts. I. The influence of set. *J. exp. Psychol.*, 1946, **36**, 71–87.
11. Weitz, J., & Flurry, Christine. A study of concept formation. *Amer. Psychologist*, 1948, **3**, 333. (Abstract)

DOUBLE ALTERNATION BEHAVIOR IN YOUNG CHILDREN

BY WALTER S. HUNTER AND SUSAN CARSON BARTLETT

Brown University

Introduction

Since Hunter's first extended report on the double alternation problem with rats (4), various additional studies have been made of the problem using rats, cats, dogs, raccoons, monkeys, and human *Ss*. Some of these investigations are listed in the references (2, 3, 4, 5, 6, 7, 8). For the purposes of the present paper, Gellermann's investigation with children is the most important.

Gellermann (2) used a T-shaped temporal maze six feet high, with outside dimensions of 18 feet by 12 feet and with alleys three feet wide. The doors which were used to regulate the pathway in double alternation were operated from outside the maze where the experimenter stood behing a one-way screen. Thirty-eight children aged three to 13 years were tested with a minimum of verbal instructions about the task. After three preliminary trials in which the *S* was familiarized with the apparatus, massed training was begun using eight responses as a trial, *rrllrrll*, and with 15 sec. between trials. The criterion of mastery was three successive trials without error. The *Ss* learned the problem in an average of 15 trials, range 4–37, except that the one three-year-old *S* and the one four-year-old *S* failed to learn in 30 and 42 trials respectively. Gellermann presents evidence indicating a strong probability that verbal (symbolic) responses control perfected double alternation in the temporal maze. After learning was completed, the *Ss* were tested on their ability to extend the series of double alternations beyond the customary eight responses. The two five-year-old *Ss* failed to extend the series. The three six-year-old *Ss* succeeded with some difficulty. Older *Ss* readily made such an extension.

The present experiment on the behavior of young children from two to seven years old seeks to determine, more nearly than did Gellermann, the age at which double alternation problems can be solved by the child. The temporal maze was not used, but rather a double alternation box apparatus similar in principle to the one used by Gellermann with monkeys (3).

Apparatus, Method, and Subjects

The apparatus used in the present experiment, Figs. 1 and 2, was similar to the alternation box-apparatus used by Gellermann (3) with monkeys. It was made of one-fourth in. plywood, painted light green. The two boxes, three in. apart, rest on the base of the apparatus. Perpendicular to the base is an upright board 24 by 18 in. from which the two boxes project. This board

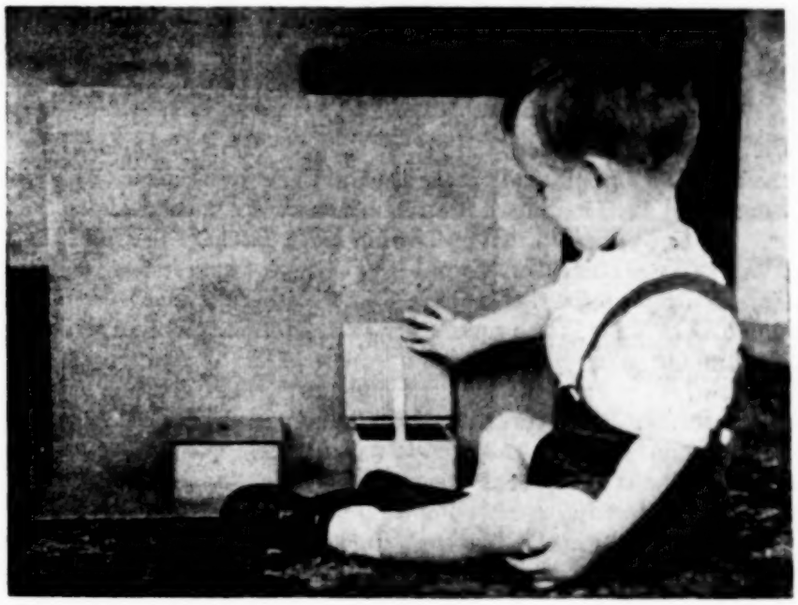

Fig. 1. The subject's side of the apparatus

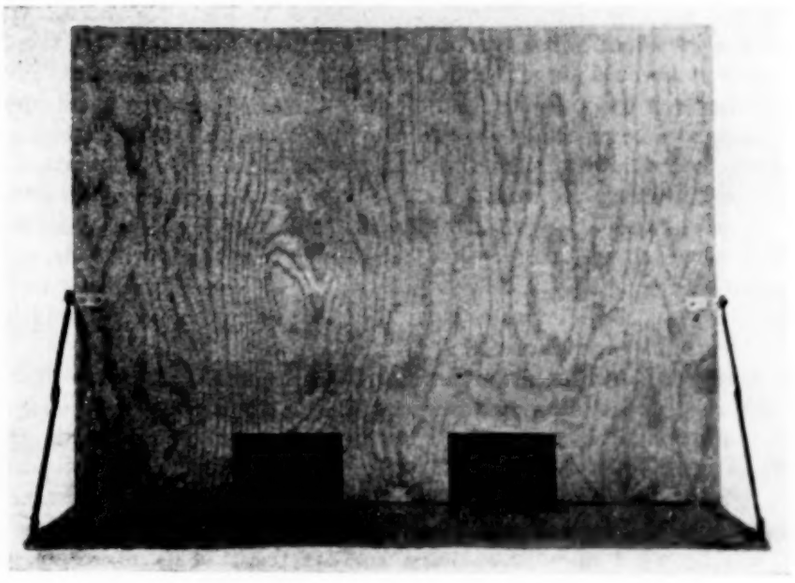

Fig. 2. The experimenter's side of the apparatus

is attached to the base by hinges so that the apparatus can be easily collapsed into a compact unit. The upright board serves as a screen to hide the manipulations of the experimenter. As the boxes rest on the base, the floor of the boxes is continuous with it. The upright screen from which the boxes project is attached to the base eight in. from the front edge of the base. Rectangular openings are cut in the upright board at the places the boxes are attached so that stimulus-

objects may be inserted by the *E* from behind the screen. The lids of the boxes project one-fourth in. beyond the box and are attached by light hinges allowing them to be opened easily by the children. Attached to each lid underneath and projecting back beyond the screen are brass strips. With these the *E* can open or close the lids. The strips serve also to indicate to the *E* behind the screen when the *S*s have lifted the lids even a fraction of an inch.

The stimulus-objects used were small pieces of candy of uniform size. However, with the young children it was necessary to use other small pieces of food that they were known to like, such as raisins and bits of cracker or marshmallow.

Thirty-one children were tested. Their chronological ages ranged from six years, nine months, to two years exactly at the time of their final experimental sessions. Their mental ages ranged from eight years to two years, three months. Nineteen children were members of a private school in Providence and the others were from professional families with whom the *E* was acquainted.

Table I lists each *S*'s chronological age and mental age at the time of his last experimental session. The mental ages were computed from scores on the 1937 Revision of the Stanford-Binet Intelligence Test. All *S*s were given this test within two months of the time of the experimental sessions. All *S*s but two were tested on Form L. The two youngest *S*s were tested on Form M.

As in Gellermann's experiment with human *S*s, verbal instructions were kept at a minimum. With *S*s four years old or more, the procedure was as follows. The experiments were always conducted in a closed room with no third person present. The apparatus was set up on the floor with *E* sitting behind it. The stimulus-objects (candy) were kept in a box behind the apparatus. When the *S* entered for the first experimental session *E* asked, "Would you like to play this game?" Then for one min. the *S* was allowed to look at the apparatus, opening and closing the boxes. Then *E* said, "I'll put candy in the boxes," and then, "Shall we begin?" The two boxes were closed and *E* put one hand into each box leaving a piece of candy in the correct one. Although *E* was sitting behind the screen her head and top of shoulders could be seen. It was for this reason that the *E* always put a hand into each box so that there would be no differential cues from sound or from *E*'s visible body movements.

The *S* opened one box or the other. If he opened the correct box, he could take out the candy and eat or keep it as he chose. If he opened the incorrect box, he found it empty. He either immediately afterward opened the correct box and got the candy or kept opening and closing the empty box until he changed to the correct one and was rewarded. At first, *E* tried to hold the lid of the incorrect box closed by means of the brass strip, but the *S*s would keep trying to get it open and would start to cry or lose all interest. This procedure was therefore abandoned and neither box was ever held shut. The *S* was always rewarded upon opening the correct box by finding a piece of candy, whether he had made an incorrect response first or not. After eight responses, *rrllrrll*, during which time the *S* had received eight pieces of candy, *E* said, "Let's stop now," and after 15 sec. said, "Let's begin again." One session usually lasted for about 12 trials, i.e., 96 responses. It varied somewhat on either side of this, depending upon the *S*'s attention. When the *S* volunteered that he wanted to stop or when he was paying poor attention, the session was brought to a close. Most of the *S*s learned in one or two sessions. An attempt was made to have the sessions as near together as possible, separated by only one or two days.

With the *S*s younger than three years old, the procedure could not be so rigid. Here, there were deviations in the length of experimental sessions; and during the first sessions with the two youngest *S*s, their mothers held them on their laps during the experiment until the children became more at ease with *E* and with the apparatus. Furthermore, there was an important deviation in procedure. Instead of breaking up the trials into eight responses, the number of responses per trial varied greatly. The *S* was given a continuous series of double alternation responses to make and there was no fixed number of responses per trial. This procedure was necessary because any interference with the *S*s after eight responses (if the *S* was still very actively opening and shutting the boxes) was met with crying, screams, or a complete lack of further cooperation. Also, some trials were shorter than eight responses when the *S* (voluntarily) crept away in the middle of a trial. Thus in order to secure the best attention and cooperation of the young *S*s, the number of responses during one session had to vary greatly. However, in calculating the number of trials in the results, four responses, i.e., one complete double alternation is considered a trial for both older and younger *S*s.

The criterion of learning for all *S*s was three correct successive series of eight responses *rrllrrll rrllrrll rrllrrll*. Although in tabulating results a trial has been considered as four responses,

eight responses were presented in a single run to the older *S*s, and three such runs were used as the criterion of success. Following the attainment of the criterion, the series was extended to 12 responses in one trial, i.e., one more sequence of *rrll*. After this, if the *S* had not by this time volunteered information indicating that he had learned the correct sequence, he was asked, "How do you manage to open the right ones first?" If the answer did not state the general solution of the problem, he was then asked, "Well, how did you learn how to do it?"

Records were kept of every response each *S* made, of all conversation of the *S* in any way pertinent to the work, and of the answers to the *E*'s questions.

Results

Table I shows the chronological age, mental age, I.Q., and number of trials necessary for each *S* to learn the problem to the criterion of three errorless successive series of eight responses: *rrllrrll*, *rrllrrll*, *rrllrrll*. In this table and in others, where the number of trials is counted for both older and younger children, a trial is considered four responses, i.e., one complete double alternation, *rrll*. The data are compiled in this way in order to keep the results of these *S*s comparable with Gellermann's results with children, where four re-

TABLE I

Chronological Ages, Mental Ages, and Number of Trials to Learn the Problem

Subject	C. A.	M. A.	I. Q.	Trials to Reach Criterion (One Trial Equals 4 Responses)
1	6-9	7-0	103	9
2	6-2	7-0	113	12
3	6-2	7-8	124	2
4	6-0	8-0	133	3
5	5-9	6-10	119	11
6	5-9	7-0	121	8
7	5-6	7-6	136	10
8	5-3	7-2	136	11
9	5-2	6-8	129	15
10	4-11	5-8	113	24
11	4-9	5-2	129	5
12	4-7	5-2	112	11
13	4-3	(?)		15
14	4-3	5-3	123	23
15	4-3	5-1	119	12
16	4-2	5-8	136	24
17	4-2	5-5	130	21
18	4-2	5-0	121	24
19	4-2	5-6	131	26
20	4-0	5-9	143	4
21	3-11	4-8	129	never in 52
22	3-9	4-7	122	12
23	3-9	4-6	129	never in 45
24	3-7	(?)		28
25	3-7	4-9	132	never in 49
26	3-4	4-7	137	never in 46
27	3-0	4-1	136	never in 43
28	3-0	(?)		never in 40
29	2-6	3-6	140	never in 106
30	2-1	3-1	148	never in 89
31	2-0	2-3	104	never in 74

sponses were counted as one trial. The youngest child who learned the problem was three years and seven months old, with an undetermined mental age.

In Table II the results are averaged for each age level in order to compare them with Gellermann's results. (The year six, for example, includes all *Ss* between six years and six years and eleven months.) Learning double alternation with the present box apparatus is evidently easier than learning the temporal maze. The average number of trials for Gellermann's six- and five-year-old *Ss* was 18.3 and 22.5 respectively as compared with 6.5 and 11 for the *Ss* in the present experiment. Also all of the present four-year-old *Ss* learned the problem in an average of 16.9 trials; and two three-

TABLE II

AVERAGE AND RANGE OF NUMBER OF TRIALS TAKEN BY *Ss* TO LEARN PROBLEM AT EACH YEAR LEVEL

Age	Subjects of Present Experiment			Gellerman's Subjects		
	Number of Subjects	Average No. Trials	Range	Number of Subjects	Average No. Trials	Range
6 yrs.	4	6.5	2–12	3	18.3	8–29
5	5	11.0	8–15	2	22.5	8–37
4	11	16.9	4–26	1	Not in 42 trials	
3	2	20.0	12–28			
3	6	Not in 40–52 trials		1	Not in 30 trials	
2	3	Not in 74–106 trials				

year-old *Ss*, in an average of 20 trials. Although these group averages decrease as age increases, there is marked overlapping in the distributions.

The correlation (rank-difference) of C.A. and number of trials to learn is .86 ± .03, using all but Subjects 13, 24, and 28. For M.A. and trials to learn, rho is .81 ± .04. In these determinations, the eight youngest *Ss* (Subject 28 omitted) who never learned the problem were each assigned an average rank of 24.5 on the number of trials to learn. (If only one *S* who did not learn is included, the C.A. versus trials correlation is .54 ± .10 and M.A. versus trials is .59 ± .10.) There was no correlation between I.Q. and trials to learn. This is not unexpected in the present group, since some of the highest I.Q.'s were possessed by children too young to learn. Gellermann found a correlation between C.A. and trials to learn of .28 ± .11 For adults his correlation with intelligence was .58 ± .09, using the Thurstone Psychological Examination for College Freshmen. Skalet (9), in a delayed reaction experiment with pre-school

children, found a correlation of .66 ± .09 between performance and C.A. and one of .75 ± .08 with M.A. (It is to be recalled that the delayed reaction test is now used at year II in Form M of the 1937 revision of the Stanford-Binet intelligence test.)

The great change in ability to solve the double alternation problem under the conditions of the present experiment, is found in the three-year age-level.[1] Of the eight three-year-old *S*s, only two learned the problem. No two-year-old succeeded. It should be noted, however, that motivation was difficult to control in the youngest *S*s. Every effort was made to give them rewards they were known to like, but with the older *S*s another motivating factor seemed to be involved. The older children were more perturbed by making incorrect responses. They often made such remarks as, "I'm going to open all the right ones," or, "You're trying to fool me," when they had made an error. These comments appear to be evidence of a desire to learn

TABLE III

CORREST SEQUENCES OF RESPONSES (*rrll*) MADE BY SOME OF THE TWO- AND THREE-YEAR OLD *S*s

Subject	C. A.	No. of Trials Before First Correct Response	No. of Double Alternation Sequences of the Following Single Response Lengths:				
			4	5	6	7	8
27	3–0	5	6	1	0	1	0
29	2–6	3	6	2	2	1	1
30	2–1	7	12	4	4	1	1
31	2–0	1	3	2	0	0	0

the correct responses although the *E* never indicated that there was a problem or system to be learned. The youngest *S*s showed relatively little of this type of behavior. They appeared to be interested in eating the candy and getting as much of it as they could; but the opening of an empty box did not seem as disturbing to them as to the older *S*s.

In the case of the children who did not learn the problem to the criterion of three correct consecutive sequences of eight responses, there were, however, some correct partial sequences of double alternation. Table III shows the number of trials before Subjects 27, 29, 30, and 31 performed the first correct double alternation series (*rrll*) and the total number of times a correct series of 4(*rrll*), 5(*rrllr*), 6(*rrllrr*), 7(*rrllrrl*), and 8(*rrllrrll*) responses was performed. Subject 30 made the largest proportional number of correct responses of these four children. Her I.Q. was 148, the highest of any *S* tested. Her M.A. was 3–1 at the time of the last experimental session when she was still younger chronologically than any *S* who learned the problem.

No one of these *Ss* was able to maintain a steady high level of performance, but the occasional series of correct responses indicate that even the youngest *Ss* had some ability to perform double alternation. The youngest *S* made no correct sequences longer than five responses. He had not yet learned to talk at the time of the last experimental session. All other *Ss* had some vocal language by the time of their last experimental sessions.

No one of the four *Ss* above referred to showed any regular decrease in percent of errors, no evidence of a learning curve. The correct runs occurred at random and were often followed by trials with a high percent of incorrect responses. It is impossible to say whether these *Ss* would have reached a consistently accurate level of performance after more trials. The delayed reaction problem seems to be a more practical test of the symbolic process at the two-year level than the double alternation problem. In delayed reaction

TABLE IV

CHILDREN'S PERFORMANCE ON THE DELAYED RESPONSE TEST
(DATA FROM MAUD A. MERRILL)

C. A. *N*	1.5 101	2 104	2.5 102	3 100	3.5 103	4 105	4.5 101	5 112
% passing 2 successes	50	80	83	90	96	99	100	—
% passing 3 successes	15	46	50	73	93	97	98	99

the stimulus is presented and withdrawn, thus setting the problem for the *S*. In double alternation the *S* must discover the problem himself. However, the high correlation of the ability to perform double alternation with age indicates that, for all but the youngest *Ss*, it is a good indication of the development of an ability which, we believe, involves a symbolic process. Both Subjects 30 and 31 were given Form M of the Stanford-Binet test. They both passed the delayed reaction test. The other *Ss* were given Form L and so were not tested on delayed reaction.

Data from the use of the delayed reaction in the Stanford-Binet examination are of interest in the present connection. Merrill[1] reports tetrachoric correlations between the delayed reaction test, 10 sec. delay, as used in the Stanford-Binet test, and the various tests on Forms L and M at the two-year and the two-year-six-months levels. These correlations range from −.040 for motor coordination to .605 for identifying objects by name. Table IV (data supplied by Merrill) shows the number of children passing the delayed reaction

[1] Personal communication from Dr. Maud A. Merrill, dated Jan. 19, 1943.

test at chronological ages from 1.5 to 5 years.[2] With respect to the behavior of the *Ss*, Merrill writes as follows: "My observation is that the child is usually intent upon the task in hand watching the examiner with laughing eagerness while the counting proceeds and ready to make his choice instantanelusly when the shield is removed revealing the boxes. The children appear to regard the test as a very delightful game and want to continue hiding the kitty."

Let us now turn to a consideration of the verbal behavior of the children in the present experiment. Although many of the *Ss* said little during the experimental sessions, some *Ss* made remarks that throw light on the learning and on the administration of self-instructions. Examples of this verbal behavior are as follows:

Subject 1. Several times said, "Oh darn it," when making an error. Just before reaching the criterion of learning she asked, "Have you put two in this one yet?"

Subject 5. Said, "Yes," after every correct response and, "No," after every incorrect response.

Subject 6. Said, "I know which to start with," preceding the fifth trial and made the first response correctly every time. After a correct response following hesitation said, "I almost opened the other."

Subject 7. Said, "Wrong one," several times after making an error.

Subject 8. Made the following remarks preceding various responses, "It'll be in here, won't it?" "I know which will be next."

Subject 9. Commented after every response. After the third response of the first trial indicated that he realized there was a problem to be learned by saying, "You're trying to fool me." Other frequent remarks were: "Right one," "Wrong one," "Got it," "I know what's next."

Subject 14. Just before learning, during a period of regular trials of *rlll*, remarked, "Two times." Later, after an error, "Let's start again. I'll get seven in six minutes."

Subject 16. "I can remember the right one sometimes." "It's hard to tell where the candy is." "Sometimes I forget which ones."

If the *S* had not volunteered the solution of the problem by the time he reached the criterion, he was asked, "How do you know which box to open?" They either spontaneously gave the solution, or responded after questioning, or said they did not know. Some examples of the three classes of responses are:

Spontaneous solutions:

"I know which to open—this (point to right) then this (R) then that (left) and then that (L) then that (R)."

"It's in there (R) two times and in there (L) two times."

"I know how now: two times and two times."

"I see how you do it; I count how many times you do it."

Answers to questioning:

"Just do—first in one box twice, then the other box twice."

"They go one and then in the same and then in the other."

"I get one (referring to the candy) in a different one after I get two in another; then I get another in the same one I've had right before the other two."

[2] Allen (1) reports for one-year-old *Ss* that 61 percent of responses at the 10 sec. delay were successful with 52 boys and 66 percent, with 48 girls.

Ss who couldn't verbalize the solution:

"I don't know—I just figure them out. It's funny."
"I don't know how I know."
"I know how 'cause I do." (Silence after further questioning.)

The records indicate that the *Ss* who were five years or older could give the principle of double alternation verbally, in fact most of them did so spontaneously. The younger *Ss*, although able to perform the problem, were not able to verbalize the solution.

All of Gellermann's *Ss* who learned the problem, mostly older than the *Ss* in this experiment, expressed the principle of 'two, two' verbally. Gellermann stated that all of his *Ss* gave evidence that extension of the series was at least partially controlled by their verbal formulation of the problem. He concluded that the verbal formulation of the solution supplements non-differential external and interoceptive stimuli to call out the double alternation of responses, i.e., that language is the symbolic process enabling human *Ss* to learn the problem.

One of the interesting questions involved in double alternation behavior is whether or not the *S* who has mastered a given series, *rrll* or *rrllrrll*, will continue to respond in a double alternation manner if he is not removed from the apparatus. Will the factor or the symbol which has controlled the original habit be of such generality that it will continue to function beyond the customary series of responses? Rats show only a slight ability to extend a double alternation series (**5, 6**). Monkeys and human *Ss*, on the other hand, show a considerable facility in this respect (**2, 3**). Gellermann (**2**) found some ability of six-year-old children to extend the series, but five-year-old children failed. In the present experiment all *Ss* who learned the double alternation problem were able to extend the series of responses whether or not they were able to formulate verbally the principle of 'two, two.' This ability to extend the series is therefore present at least in some children as early as three years and seven months of age.

Summary

1. Thirty-one children ranging in chronological age from two years to six years and nine months were trained on a double alternation box-apparatus. The youngest child to attain the criterion of three successive errorless trials of eight responses each was three years and seven months old. Younger children could perform double alternation, but at a lower criterial level.

2. There was a positive correlation between trials to learn and both C.A. and M.A.

3. Children five years old or older could state verbally the principle of double alternation by the time they had reached the criterion of learning. Younger children could not do so, even though they had mastered the problem.

4. Children at least as young as three years and seven months can extend their double alternation responses beyond the length of the series on which they have been trained.

5. The double alternation problem is not as well adapted to the behavior characteristics of the young child as is the delayed reaction test, although both appear to involve in their mastery a symbolic process. In the delayed reaction, the disappearance of the stimulus object sets the problem for the *S*, whereas in the double alternation problem the *S* must discover for himself the pattern of response required.

(Manuscript received October 22, 1947)

References

1. Allen, C. N. Individual differences in delayed reaction of infants. *Arch. Psychol.*, 1931, **19**, No. 127.
2. Gellermann, L. W. The double alternation problem: II. The behavior of children and human adults in a double alternation temporal maze. *J. genet. Psychol.*, 1931, **39**, 197–226.
3. Gellermann, L. W. The double alternation problem: III. The behavior of monkeys in a double alternation box-apparatus. *J. genet. Psychol.*, 1931, **39**, 359–392.
4. Hunter, W. S. The temporal maze and kinesthetic sensory processes in the white rat. *Psychobiol.*, 1920, **2**, 1–18.
5. Hunter, W. S., & Hall, B. E. Double alternation behavior of the white rat in a spatial maze. *J. comp. Psychol.*, 1941, **32**, 253–266.
6. Hunter, W. S., & Nagge, J. W. The white rat and the double alternation temporal maze. *J. genet. Psychol.*, 1931, **39**, 303–319.
7. Keller, F. S. A new type of double alternation. *J. genet. Psychol.*, 1937, **51**, 454–459.
8. Schlosberg, H., & Katz, A. Double alternation lever-pressing in the white rat. *Amer. J. Psychol.*, 1943, **56**, 274–282.
9. Skalet, M. The significance of delayed reactions in young children. *Comp. Psychol. Monogr.*, 1931, **7**, No. 4.

THE EFFECT ON GROUP PERFORMANCE OF AN INDIFFERENT AND NEGLECTFUL ATTITUDE SHOWN BY ONE GROUP MEMBER

BY DAVID ROSENTHAL

Washington School of Psychiatry

AND

CHARLES N. COFER

University of Maryland [1]

Most writers agree that the term 'morale' in some way involves motivation toward group goals (cf. **1, 2, 3, 6**), which function in some relation to goals that an individual holds for himself. Despite this emphasis, relatively little effort has been devoted to studying the process of group goal formation or to the effect on these goals of experimentally introduced variables.

The present study was designed to investigate, in the context of group activity toward a group goal, some of the relationships between individual and group goals, and the effects on both of an attitude of 'indifference and neglect' manifested by one member of the group. It is hoped that by experimental investigation of group goal processes further analysis of the morale concept may become possible.

METHOD

General plan

The task employed was that of dart-throwing. The *S*s were male students enrolled in introductory psychology courses. They were told that they were being compared directly with a group of federal government employees, as a part of a large project the purpose of which was to compare intellectual and motor processes of college students with groups of different academic, occupational and social status. The task was described to the *S*s as a group project in which individual scores were unimportant but in which the group score they obtained was of primary significance. They were urged to assist one another to improve their techniques of dart-throwing. The *S*s were told that the group of employees with whom they were being compared had achieved a certain *group* score on the task and that they were to try to equal or surpass it. Actually, the score set was a fictitious one, so high that it was not reached by any group used in this study.

Each *S* made five successive dart throws in each of nine trials, and his score was reported as the sum of his scores on the five throws. The group score was reported as the sum of the scores for all of the *S*s in a group for a trial. The target was of the usual variety and was graded to give scores from zero (outside the ring) to seven (bull's eye).

Groups of four and five *S*s were employed. In the experimental groups, the 'Plant' (see below) was one member of each group. There were seven control and seven experimental groups

[1] This study was conducted at George Washington University, Washington, D. C. The present paper is an abridgement of a Master's thesis, prepared by Rosenthal.

of four *Ss* each, and five control and five experimental groups of five *Ss* each. Thus, there were 53 control *Ss* and 41 experimental *Ss* (not counting the 'Plant,' who participated in each of the 12 experimental groups).

'*Attitude scales*'

To gain an index of the attitudes or the 'motivation' of the *Ss*, five scales were devised, in accordance with Likert's technique (5), which was modified as necessitated by the demands of this study. The scales, in the form in which they were administered to the *Ss*, are presented below. (The material here presented in parentheses did not appear on the blanks given to the *Ss*.) They were administered to all groups, following the first and ninth trials. The five items on each scale were scored from plus two (favorable attitude) through zero (neutral or indifference point) to minus two (unfavorable attitude). Since the scales were devised solely for the purposes of this study, in which the performances of control and experimental groups on them were simply to be contrasted, their reliability and validity may be seen from the results of the experiment. It should be mentioned that differences with respect to these scales between the control and experimental *Ss* do not necessarily validate them as measures of motivation.

Attitude scales used in this study

Below are listed 5 questions. Under each question are listed 5 statements. Answer *all* questions by checking the one statement under each which best describes how you feel. Be as frank and as accurate as you possibly can. Return the answered form to the experimenter.

Scale I

Do you think that experiments of this sort are of value to science?

.....Very worthwhile
.....Quite definitely not worthwhile
.....Probably not worthwhile
.....Probably worthwhile
.....Don't know

Scale II

How do you personally feel about this experiment?

.....Am not enjoying it very much
.....Am enjoying it very much
.....Am indifferent to it
.....Am enjoying it
.....Am not enjoying it at all

Scale III

What do you feel is the probability that this group will achieve the standard?

.....It probably will
.....Don't know
.....It will not
.....It will
.....It probably will not

Scale IV

Do you think the other members of the group will try their utmost to achieve the standard?

.....Don't know
.....Will not try very hard
.....Will try their utmost
.....Will not try at all
.....Will try fairly hard

Scale V

How do you feel about trying to help the group reach the standard?

.....Will try hardly at all
.....Will keep on trying fairly hard
.....Don't expect to try very hard
.....Don't know
.....Will keep on trying my hardest

Level of aspiration forms (LA forms)

Although the precise significance of the level of aspiration (4) is not settled, it is usually agreed that it involves motivation in some manner. It was, therefore, used in this study. LA forms were filled in by all *Ss* following each trial. These forms contained the following questions:

(1) What score did you make last time? (2) What score do you expect to make next time? (3) What score did the group make last time? (4) What score do you expect the group to make next time? The response to question 2 is here considered to be the individual's level of aspiration for himself. The response to question 4 is here considered to be the individual's level of aspiration for the group of which he is a member. So far as we can determine, this is the first use of a group level of aspiration (question 4).

Vocal, collective level of aspiration

After the third and seventh trials and following the execution of the LA forms, the groups were asked, "What score would you accept next time as a mark of satisfactory progress toward reaching the goal?" The group was asked through discussion to agree, if possible, on an acceptable score.

The Plant [2]

The control and experimental groups were treated identically through the first trial. After this trial, one *S*, who served in all of the experimental groups, began to behave in a manner designed to be indicative of 'indifference and neglect.' Except for his behavior and its effects, the control and experimental *S*s were treated similarly throughout the experiment. The Plant (P), a senior majoring in psychology, collaborated with *E*. Prior to the experiment, he was carefully instructed concerning the nature of the experiment and his role in it. A series of acts and statements to be employed by P in the experiment was worked out; they were chosen because they were acceptable to P's habitual mode of expression and because they were believed to embody the concepts of indifference and neglect.

The Plant's behavior of indifference and neglect

Beginning with the second trial, P started to act with indifference and neglect toward the goal and the experiment itself. His behavior was essentially negativistic and was manifested broadly as a characteristic attitude. He showed a complete lack of interest in the experiment, in his own performance-caliber, and in the progress of the group. At no point did he behave positively toward the goal. He manifested indifference throughout the experimental session, participated haphazardly and even begrudgingly. His manner was obviously slipshod and undirected, accompanied by appropriate shoulder-shrugging and facial expression. In addition to the broad attitude-pattern he manifested, he also did and said the following:

1. When about to throw in the second trial, he said, "This sort of stuff is pretty silly, isn't it?"
2. When it was P's turn to throw in the third trial, he went to the sink to take a drink of water and showed no concern about delaying the trial.
3. When the group was trying to agree on an acceptable mark of progress after the third trial, P said, "I can't see what difference it will make one way or the other."
4. When about to throw in the fourth trial, P said, "I don't get the idea. What can they find out that they don't already know?"
5. When filling out his LA form after the fourth trial, P called to the *E*, "I didn't hear you. What score did we make the last time?"
6. During the fifth trial, P seemed to become interested in a string apparatus, which was in the laboratory, plucking the strings and ignoring everything going on about him.
7. During the sixth trial, P said, "What a waste of time this turned out to be."
8. During the seventh trial, P seemed to become interested in a bell apparatus, which was in the laboratory, indulging himself in some bell-ringing.
9. After trial seven, when *E* was seeking an acceptable mark of progress from the group, P said, "I still can't see what difference it will make."
10. In trial eight, P took a second drink of water, as in trial three.
11. In trial nine, P said, "I'm getting tired of this. Let's get it over quick." He then threw his darts as rapidly as possible.

[2] Mr. Richard Kelly Greenbank served as the Plant. We are immeasurably indebted to Mr. Greenbank, whose cooperativeness and effective acting really made this study possible.

12. To avoid the suggestion of complicity with *E*, P was permitted to do and say other things which came to him spontaneously and which fitted the situation within the group at the moment. When one *S* threw well, P exclaimed, "I bet they give him a Bachelor of Darts Degree." He also produced several other remarks and mannerisms which seemed to fit in well with his role in the group at the time. P was not instructed to throw poorly since the effect of 'indifference and neglect,' not poor performance, was the factor under investigation. P was told to try to obtain a score which would fall somewhere in the middle of the group. Of course, this was not always possible. At times it was difficult for P, even after considerable practice, to obtain high scores while trying to throw in a manner suggestive of indifference.

Results and Discussion

Attitude scale scores

The essential data for the rating scales are shown in Table I. For each individual scale, the possible scores ranged from +2 to −2 and for the five scales summed together ranged from +10 to −10.

TABLE I

Average Initial and Terminal Scores on the Attitude Scales for the Control and Experimental Groups

Scale		Ave. Initial Score	Ave. Terminal Score	Difference	Difference Between Groups	Critical Ratio
I. Value to science	Control	1.15	1.25	0.10	0.12	not significant
	Experimental	1.12	1.10	−0.02		
II. Personal enjoyment	Control	1.23	1.49	0.26	0.10	not significant
	Experimental	.93	1.20	0.27		
III. Belief in goal attainability	Control	.62	0.43	−0.19	0.96	4.0
	Experimental	.44	−0.71	−1.15		
IV. Evaluation of group effort	Control	1.36	1.70	0.34	1.34	6.3
	Experimental	1.44	0.44	−1.00		
V. Personal effort to be expended	Control	1.83	1.87	0.04	0.00	not significant
	Experimental	1.76	1.80	0.04		
Total	Control	6.17	6.74	0.57	2.42	5.8
	SD	1.74	2.14			
Total	Experimental	5.67	3.82	−1.85		
	SD	2.09	1.86			

Some differences in scale scores between the control and experimental groups appeared on the initial administration of the five scales (following trial 1). The differences, however, are not statistically significant either for the scales considered individually or considered as a whole.

Table I likewise shows the individual and total scale scores for the second administration (following trial 9). Here, it is seen that the total scale scores of the control groups increased during the experiment by 0.57, whereas the scale scores in the experimental groups decreased by 1.85 points. The total difference between the final scale scores of the two groups is 2.42; this difference is statistically significant, yielding a critical ratio of 5.8.

Further inspection of Table I reveals that only two of the five scales contributed significantly to the total difference just discussed. The most significant difference is in terms of Scale IV, in which the respondent was asked to evaluate the effort to be expended by the other group members in attaining the goal. Scale III, involving the evaluation of goal attainability, shows a decline for both groups, but the decline is significantly larger in the experimental groups than in the control groups.

Two factors, other than the indifference and neglect shown by P, might account for these differences. One is the actual dart-throwing performance of the two groups. Study of trial by trial performance of the two groups, however, shows no consistent superiority for either group and in six of the nine trials the difference between the group averages was less than one point per man (Table III, below). Clearly, the performance caliber of the groups was approximately the same.

The second factor is P's performance. He had been instructed to try for an average actual performance. That he did so is indicated by the fact that the mean score per man per trial for the experimental groups is essentially the same, whether the groups contain P or not.

It is clear, then, that the experimental group suffered a greater decline of reported belief in goal attainability than did the control group with no more objective (performance) reason to do so. It further showed a development of a reported disbelief that the other group members would try hard to attain the goal, again without justification from the standpoint of performance. It seems reasonable to conclude that P's expression of an attitude of indifference and neglect was the source of these changes in the groups of which he was a member. (In one experimental group, an *S* remarked that performance would be better if some people weren't indifferent about it.)

It is noteworthy that no significant differences between the groups appeared for the other scales. Personal enjoyment of the task was reported as increasing in about the same amount by the groups. Perhaps the shortness of the session and the intrinsic interest of dart-throwing account for this; longer sessions with less enjoyable tasks, especially if they required group cooperation rather than individual performances, might have shown differences between the groups on

this scale. The groups did not differ significantly in their estimates of the significance of the experiment to science.

The failure of Scale V to show differences between the groups raises an interesting problem. The experimental groups had suffered a severe decline in reported belief in goal attainability and reported that they believed that the other group members would not try to reach the goal. The control group also showed some decline in reported belief in goal attainability. Despite these factors, the individuals in both groups maintained that they would still personally work as hard or harder than before to attain the goal. For the control group this is not surprising, since they expected the other group members to do likewise (Scale IV). For the experimental group, at least two explanations are possible: (1) that the experimental *Ss*, having given up hope in goal attainability (Scale III) and belief in the efforts of the others (Scale IV), had substituted personal goals for

TABLE II

COMPARISON OF TOTAL ATTITUDE SCALE SCORES OF LARGE AND SMALL GROUPS

	Control Groups		Experimental Grousp*	
	Large	Small	Large	Small
Mean, initial administration of scale (total)	6.32	6.04	5.1	6.24
Mean, terminal administration of scale (total)	7.04	6.47	3.05	4.57
Mean difference	.72	.43	−2.05	−1.67
N	25	28	20	21

* P's attitude scale scores were excluded from these averages.

the group goal and were continuing their efforts because of these personal goals; (2) that the experimental *Ss* stated that they would continue to try hard, because they believed it was the thing to say, or that it was expected of them. This factor may well have operated, although the *Ss* had been encouraged to express themselves very frankly. It is difficult to see why this factor should have operated in Scale V when it apparently did not operate in Scales III and IV.

Group size and attitude scales

Data were available from the study which allow a preliminary analysis of the following question: Does the indifferent and neglectful behavior of one group member have different effects dependent upon the size of the group? This was not the major problem of the study, but groups of four members and five members (both including P) were used. Table II shows the relevant data for the scales.

In the control groups the *increase* in attitude scale scores was about two-thirds greater for the large groups than for the small groups. In the experimental groups the opposite reaction occurred, the larger groups showing a greater *deterioration* (about 23 percent) than the smaller. The difference in deterioration between the two sizes of experimental groups is, however, not statistically significant, its critical ratio being .66. The small *N* and the low mean obtained on the initial administration of the scales by the large experimental groups introduce the possibility of sampling errors. Nevertheless, the data do seem to point in the direction of a difference in the degree of motivational decline between the groups of different size and this problem might well be investigated further.

Level of aspiration

Table III presents the following data: (1) The mean score in dart-throwing per man per trial; the data for the experimental *S*s include P's scores; (2) predictions of group performance trial by trial (group LA) presented as average predicted score per man (P's 'predictions' were omitted from these computations); (3) predictions of own individual performance trial by trial (individual LA), presented as average predicted score per man (P's 'predictions' were also omitted here).

TABLE III

MEAN PERFORMANCE AND LEVEL OF ASPIRATION SCORES PER TRIAL FOR THE CONTROL AND EXPERIMENTAL GROUPS

Data for the Experimental group include P's performance but do not include his level of aspiration scores.

Trial	Performance		Group LA		Individual LA	
	Control	Experimental	Control	Experimental	Control	Experimental
1	15.9	16.1				
2	18.6	18.8	20.0	19.4	22.0	19.9
3	20.0	19.8	21.3	21.1	23.8	22.4
4	21.2	19.0	21.8	21.9	24.2	24.5
5	20.0	21.1	22.8	21.5	25.5	24.7
6	22.5	21.4	22.4	22.7	25.1	25.4
7	22.6	22.2	24.0	23.1	26.0	25.4
8	22.1	21.3	24.2	23.4	26.3	25.7
9	21.8	21.2	24.4	22.8	26.4	25.7
10			24.3	22.9	26.6	25.7
N	53	53	53	41	53	41

In both experimental and control *S*s, predictions for one's own performance were consistently higher than predictions for other group members. The group LA for both groups of *S*s seems to vary with success and failure and to become stabilized within narrow limits during the last few trials at a level slightly above that of actual per-

formance. Performance was likewise fairly stable in the last few trials.

Thus the group predictions for the last four trials in the control groups vary from 24.0 to 24.4, and performance in the last four performances varies from 21.8 to 22.6. In the experimental *Ss*, the last five predictions range from 22.7 to 23.4, whereas the last five performance trials yield scores varying from 21.1 to 22.2. The group LA is somewhat higher in the controls than in the experimental *Ss* for these terminal trials, but the control group's performance in these trials was likewise slightly better than that of the experimental *Ss*.

Essentially the same statements hold for the individual LA for both groups. Discrepancy scores were also compared. These scores are the difference between the performance just completed and the prediction of score on the trial immediately to follow. As may be seen by study of Table III, the discrepancy scores of the control and experimental groups do not differ markedly, and neither group was consistently above or below the other in size of discrepancy score throughout the trials.

Although the LA findings reported do not clearly testify to the motivational changes as revealed in the rating scales, it is noteworthy that in this situation the individual LA is found consistently to be higher than the group LA and that the course of the two measures is highly similar. It may well be that in making the two predictions, the *Ss* really did not differentiate between them in this situation, except in terms of a relatively constant quantitative difference. The highly complex nature of the level of aspiration may be responsible for the failure of this measure as used here to show the motivational changes clearly apparent in the other data. Both groups made predictions that were fairly close to actual performance, and it may be that the probability of success (4) as seen by the *Ss* seemed so low that it obscured the functioning of motivational factors.

Vocal, collective LA

It will be recalled that another form of group LA was obtained through discussion after the third and seventh trials. The groups were asked to agree on a score that they felt would be a mark of satisfactory progress toward the goal, but agreement was not required and hence the aspirations as stated cannot be compared. Rather, ease of agreement was the characteristic it was desired to observe, because it was supposed that the better motivated groups might reach agreement more easily than would less well motivated groups.

Agreement in reaching a group level of aspiration through discussion was divided into four categories:

1. *Immediate agreement*—where all members spontaneously gave or agreed to the same score.

2. *Quick agreement*—where after brief discussion all agreed to an acceptable mark of progress.

3. *Fair agreement*—where agreement was reached after prolonged discussion; or where discussion was brief and stopped before agreement was reached but where the range of scores given was less than 10.

4. *No agreement*—where after much or little discussion a range of scores greater than 10 was given and where the different scores offered were left unreconciled.

Obviously, the experimental groups had much more difficulty in coming to agreement and this difficulty increased considerably from the third to the seventh trials (Table IV). This inability to agree

TABLE IV

NUMBER OF SUBGROUPS IN WHICH THE VARIOUS DEGREES OF AGREEMENT OCCURRED ON VOCAL, COLLECTIVE LEVELS OF ASPIRATION

	Third Trial		Seventh Trial	
	Control	Experimental	Control	Experimental
Immediate agreement	2	1	3	0
Quick agreement	8	4	8	1
Fair agreement	1	7	1	4
No agreement	1	0	0	7
Number of subgroups	12	12	12	12

can presumably be attributed to the disruptive influence of P's behavior. Thus, indifference and neglect not only provoke deterioration of motivation toward group goals but also provoke a disintegration of commonality of group thought and feeling.

The reader is reminded that during the taking of each vocal, collective level of aspiration the Plant said only that he "couldn't see what difference it makes." He was not permitted to state any score but was required to agree with any score agreed upon by the other group members. This procedure was followed in each group.

SUMMARY

The problem of this study was to determine if non-participative behavior on the part of one group member, such as the manifestation of an attitude of 'indifference and neglect,' would produce measurable effects on attitudes toward the group goal of other group members and on both their individual levels of aspiration and their aspiration levels for the performance of the group. Male college students, as-

signed at random to the experimental and control subgroups, were used as Ss.

It was found that, if a group member behaved with indifference and neglect, the attitudes of the group members shifted significantly in the direction of disbelief in goal attainability and of disbelief that other group members would whole-heartedly participate to achieve the goal. It is suggested, but not demonstrated, that the experimental group members tended to give up the group goals and to substitute for them personal goals. Group size was seen to have a possible relationship to the effectiveness of the indifferent S's behavior. Ease of agreement in setting a group level of aspiration through discussion was found to be greater in the control groups than in the experimental groups. Other level of aspiration data, as obtained here, did not significantly differentiate between the groups. A consistent difference was found between the size of predicted performance for the group and predicted performance for the individual when both predictions were made by individuals without consultation with others. The former prediction, designated here the group level of aspiration, seemed to follow the same course and to be affected by success and failure in ways similar to the individual level of aspiration.

(Manuscript received September 22, 1947)

References

1. Allport, G. W., & Veltfort, H. R. Social psychology and the civilian war effort. *J. soc. Psychol.*, 1943, **18**, 165–223.
2. Blankenship, A. B. Methods of measuring industrial morale. In: Hartmann, G. W., & Newcomb, T. W. (eds.), *Industrial conflict: a psychological interpretation.* New York: Cordon, 1939. Pp. xi + 583.
3. Child, I. L. Morale: a bibliographical review. *Psychol. Bull.*, 1941, **38**, 393–420.
4. Lewin, K., *et al.* Level of aspiration. In: Hunt, J. McV. (ed.), *Personality and the behavior disorders.* New York: Ronald Press, 1944. Vol. 1.
5. Likert, R. A technique for the measurement of attitudes. *Arch. Psychol.*, New York, 1932, **22**, No. 140.
6. Schmeidler, G. R., & Allport, G. W. Social psychology and the civilian war effort, May 1943–May, 1944. *J. soc. Psychol.*, 1944, **20**, 145–180.

THE RELATIVE SWEETNESS OF SUGARS: SUCROSE AND DEXTROSE *

BY P. E. LICHTENSTEIN

Indiana University

INTRODUCTION

The problem of relative sweetness has been attacked by various psychophysical methods. Each of these, however, possesses certain defects which may hinder its effective application. In general the methods may be summed up under two headings: (a) the (absolute) threshold method, and (b) the method of successive comparison. Method *a* has, for various reasons which will become clear in the following discussion, been most frequently applied by investigators in relative sweetness studies.

Using the threshold method Biester, Wood and Wahlin (1) found the lowest concentration of a sugar solution which was perceptibly sweet to a number of subjects. By determining sweetness thresholds for a number of sugars these investigators were able to obtain numerical sweetness ratings. The sweetness of any sugar is taken as the reciprocal of its threshold concentration. The figure thus obtained is then adjusted to sucrose as an arbitrary standard equal to 100 (3). In practice it is simply necessary to divide the sucrose threshold concentration by the threshold concentration of the comparison stimulus and multiply by 100. Numerical ratings for several sweet substances have also been obtained in this manner by Carr, Beck, and Krantz (2) and Willaman (6).

When method *b* is used relative sweetness values may be obtained by dividing the concentration of the sucrose solution used as a standard by the concentration of a solution of the comparison stimulus equivalent in sweetness to the standard. The resultant figure is then multiplied by 100. One should note that relative sweetness is not a term which is operationally equivalent for the two methods. The lack of clarity which had characterized the term relative sweetness should serve to caution workers in this field. If relative sweetness is defined in terms of threshold determinations, one may not quarrel with the definition. On the other hand, the restriction of the term to threshold measurements does not appear particularly fruitful.

The threshold method has the advantage of allowing for rapid determinations since only one point, the threshold, is experimentally determined for each sweet substance. In addition the method avoids the necessity of employing successive comparison which involves responding after a delay to

* This investigation was carried out at the University of Massachusetts under the direction of Dr. C. C. Neet.

one of the comparison stimuli. On the other hand, the use of the method of limits to obtain a threshold has several shortcomings. Concerning these one might point out that thresholds tend to be unstable and difficult to determine with accuracy since they vary from one individual to another and in the same individual from time to time. Further, and perhaps of greater importance, is the fact that the comparison of thresholds can give no reliable information concerning the relative sweetness of supraliminal stimuli. Consequently results based solely upon threshold determinations become misleading when taken to apply over the entire range of stimulus concentrations. One might determine a differential threshold both for sucrose and for any comparison stimulus and use the result as a basis for extrapolation. Such an extrapolation, since it would assume the validity of Weber's Law, appears to be gratuitous and perhaps unjustified. It might also be pointed out that the threshold method does not actually involve the comparison of stimuli. Although the lack of such comparison may not have undesirable consequences, a direct comparison might be more desirable.

Since simultaneous comparison of stimuli is not feasible in taste study, a direct comparison of stimuli entails the use of the method of successive comparison. Biester, Wood, and Wahlin (1) tested the method of successive comparison and took their results to indicate the unreliability of this procedure. They found, for example, that when *S* compares a stimulus in a graded series with one given just prior to it the interval which separates the stimuli must be large if the difference is always to be detected. An interval of 1.5 percent in a sucrose series ranging from 0.75 percent to 9.75 percent gave 94.4 percent accurate judgments. In a galactose series, however, an interval of 4.75 percent was necessary to give 96.2 percent accurate judgments. The authors state that, "In the galactose series three decisions indicated that one solution was as sweet as another having a concentration of 4.75 percent less" (1, p. 390). It is on the basis of such evidence that the method of successive comparison was rejected.

The above conclusion can only be drawn by failure to take into account the differential threshold. Since galactose is only slightly sweet, a small j.n.d. could scarcely be expected. The differential threshold probably assumes considerable importance for studies of relative sweetness. Its significance should not be ignored, and least of all should the existence of a large j.n.d. be so interpreted as to provide a basis for rejecting any method involving successive comparison. Another aspect of the findings of Biester, Wood and Wahlin on the method of successive comparison concerns the error of measurement. Psychophysical measurement like any other measurement shows experimental error. Such error, involving the variability of response, is not eliminated either by using simultaneous comparison or the threshold method.

Willaman (6), using the threshold method, found the sweetness of levulose to be 173 whereas earlier investigations using successive comparison had led to the lower figure of 108. Willaman implies that the discrepancy in results is due to the inaccuracy of tests involving memory. As we have already seen there is no good evidence for such an assumption. Neither is

there any a priori reason for assuming that the relative sweetness of a substance can be represented by a single numerical value. Since the threshold method and the method of successive comparison usually involve very different concentrations of the sweet substances, there seems to be no reason for rejecting one sweetness value because of its lack of agreement with another value obtained by another method but at a different concentration. Sweetness values might well be found to vary as a function of the concentration.

Our theoretical argument against the use of taste thresholds and the subsequent attempt to represent relative sweetness by a single numerical value finds support in the work of Renner (5). This investigator found that the relative sweetness of sucrose, glucose and fructose changes with the concentration. He states that the sweetness of glucose is 53 at a concentration of eight percent (sucrose = 100) but rises to 88 at 35 percent. Obviously no single numerical value can adequately represent the relative sweetness of glucose in this instance.

In the light of the above criticisms it appeared justifiable to develop a method adapted to the problem of comparing sweet stimuli regardless of the concentration. Such a method must almost necessarily be a variation of the method of successive comparison. One purpose of the present investigation was to show that the method of successive comparison may be modified to yield reliable results in the determination of relative sweetness. A second purpose was to establish relative sweetness values for two sugars, sucrose and dextrose, over a wide range of concentrations.

Results previously obtained concerning the relative sweetness of dextrose show little agreement. The interpretation of results is complicated by the failure of investigators to use methods which could be expected to render comparable findings or in some cases of authors to specify clearly the method used or the concentrations compared. In other cases investigators have failed to specify the purity of the sugars used. Sometimes one finds dextrose and glucose used as synonymous terms for the relatively pure product dextroglucose while at other times glucose is used to refer to commercial glucose, a mixture of dextroglucose, maltose and dextrins. Such facts explain in part how one investigator arrives at a sweetness value for dextrose of 45 while another reports a value as high as 80.[1]

Subjects, Materials, and Procedure

Five *Ss* were used in the experimental investigation, two male and three female. All were undergraduate college students. Pure sucrose conforming to U.S.P. × 1 was used in all tests. This product contains 99.9 percent sucrose and is 68.1 percent soluble in water at 30° C. Sucrose, or ordinary cane sugar, has apparently a pure sweet taste and it has been used in previous experiments as a standard for the comparison of other sweet substances.

The other sugar used was anhydrous dextrose which is at least 99.5 percent dextrose and

[1] Further references and discussion are given in Biester, Wood and Wahlin (1).

contains not more than 0.5 percent moisture. This conforms to the specifications of the U.S.P. × 1. Dextrose is also known as glucose, grape sugar, or corn sugar. It is slightly less soluble than sucrose, being 54.6 percent soluble in water at 30° C., and in solution it is slightly more fluid than sucrose solutions of equal density. Many higher sugars yield dextrose upon hydrolysis, and dextrose is the form in which nearly all sugar is utilized by the human body.

Soultions of the two sugars were prepared with distilled water and the percentage concentration was determined by the specific gravity method. Occasionally solutions were checked with a refractometer and no discrepancies were noted. Solutions were kept in clean stoppered flasks which were appropriately labeled. New solutions were made up every third day to guard against the formation of molds and subsequent acid production.

Each *S* sat at a small table upon which were placed a 300 cc. beaker containing distilled water for rinsing the mouth and a 600 cc. beaker into which he could expectorate. *S*s were blindfolded and the nostrils stopped with absorbent cotton. Filter paper cut into strips measuring about one and one-half in. square was used to remove visible moisture from the tongue before each stimulus presentation.

Solutions were placed upon *S*'s tongue with a medicine dropper. The medicine droppers were immersed in distilled water when not in use and were frequently washed with hot water. A possible difference in the size of the drops was not thought to be so great as to produce any appreciable error.

A modification of the method of limits was used in the experimental investigation. A standard sucrose solution was compared with a series of dextrose solutions which ranged from definitely less sweet to definitely sweeter than the standard. Ascending and descending runs were made. In an ascending run *S* compared the least sweet dextrose solution with the standard sucrose and proceeded in discrete steps to the point at which the sucrose and dextrose appeared to be of equal sweetness. In a descending run *S* started with a dextrose solution clearly sweeter than the standard and proceeded down, step by step, until an equality point was reached. Judgments of 'sweeter,' 'less sweet' and 'equal' were allowed. In any series the transition from a judgment of 'less sweet' to 'equal' or 'sweeter,' or from 'sweeter' to 'equal' or 'less sweet' established the equality point for that run. Thus, although three response categories were used, *E* treated the data as though there were only two. Judgments of 'doubtful' necessitated that the trial be repeated. To guard against the possibility of *S*'s responding to the length of the run, *S* was started at different points in the series of dextrose solutions in the various runs. The time error was controlled by presenting the standard first in one-half of the trials and the comparison stimulus first in the other half. The order of presentation was varied in random fashion.

In a psychophysical study of this sort atypical judgments sometimes occur. That is, *S* may call a dextrose solution sweeter than or equal to the standard even though in most judgments he calls it less sweet. In such instances a repetition of the trial was made. If the first atypical response was reversed, the run continued in the normal manner. If, however, the first judgment was repeated, it was allowed to stand as the equality measure for that run.

The following instrustions were given to each *S*:

1. Your task is to compare two sensations of sweetness with regard to intensity. The second of a pair of stimuli will always be compared with the first.

2. When the solution has been placed on the tongue close your mouth and press the tongue forward against the roof of the mouth close to the front teeth. Be as consistent as possible in the execution of this procedure.

3. After your mouth has been rinsed and your tongue dried the first stimulus will be given and you will try to form as distinct an image of the sensation as possible.

4. At the command 'Ready' you will rinse your mouth with water from the beaker held in your right hand. At the next command 'Ready' you will expectorate into the large beaker held in your left hand.

5. Your tongue will again be dried and the second stimulus of the pair will be given. At the command 'Report' you will signify your judgment by tapping on the table with the beaker held in the right hand. If the second stimulus is less sweet than the first, tap once; if the second is sweeter than the first, tap twice; if the two are equal, tap three times. Failure to respond will indicate that you are doubtful.

6. After reporting rinse the mouth without command and expectorate at the command 'Ready.'

7. Do not talk at any time except to answer questions from the experimenter or to ask necessary questions concerning your procedure.

Two separate experiments were carried out. Experiment 1 consisted of three measurements for each *S* relating dextrose to the standard sucrose. The three parts of Experiment 1 are referred to as Series 1, Series 2, and Series 3. In Series 1 a series of dextrose solutions was compared with a standard 10 percent sucrose solution. In Series 2 the sucrose standard was 25 percent and in Series 3 it was 40 percent.[2]

Experiment 2 was designed to determine the sweetness of a sucrose-dextrose mixture by comparing a series of concentrations of the mixture with a standard sucrose solution. This experiment was carried out at only one concentration of the standard with a mixture two-thirds sucrose and one-third dextrose. Results in Experiment 2 were compared with a calculated sweetness value of the mixture based upon the separate sweetness values of sucrose and dextrose

Results—Qualitative

Ss reported few qualitative differences in the sugars at low concentrations. In Series 3, however, qualitative changes in the taste of dextrose became apparent. All *Ss* reported a slightly unpleasant bitterness in dextrose solutions of high concentration at one time or another. The bitterness accompanied an intense sweetness. Sucrose at high concentration showed no similar effects, *Ss* describing it as purely sweet. At high concentrations sucrose was uniformly described as pleasant and dextrose as unpleasant. One *S* described dextrose as 'digging into' a given area of the tongue in contrast with the greater spread of sucrose over the tongue. Other investigators have noted qualitative effects with change in the concentration, observing that strong dextrose solutions give rise to burning sensations in the throat (4).

Undoubtedly qualitative changes in dextrose, correlated with a change in concentration, introduce a complicating factor in the experiment which cannot be entirely overlooked.

Results—Quantitative

Experiment 1

Series 1.—The comparison of a 10 percent sucrose solution with a series of dextrose solutions yielded the results shown in Table I.

TABLE I

Dextrose Concentration Equivalent to 10 Percent Sucrose and Corresponding Sweetness Value of Dextrose

Subjects	Mean (conc. dextrose)	SD	SE_M	Sweetness Value (dextrose)
MS	15.4	2.12	.27	65
MC	15.6	1.72	.22	64
MM	15.3	1.34	.17	65
GS	15.6	1.37	.18	64
GF	15.4	2.14	.27	65
All *Ss*	15.5			65

[2] We have followed conventional practice in sugar chemistry of specifying concentration as percent by weight, that is the percent which the weight of sugar is of the total weight of solution.

Dextrose concentrations for the first series as obtained by refractometer readings were 9.5, 10.3, 11.1, 12.0, 12.6, 13.2, 13.8, 14.5, 15.5, 16.0, 16.8, 17.5, 18.1, 18.5, 19.5, 20.0, and 21 percent.

Sweetness values for dextrose are seen to be practically the same for all *Ss*. The fact that SD values vary considerably for different *Ss* may be taken to mean that discrimination is much keener in one *S* than another or that different *Ss* use slightly different methods of judging.

Series 2.—A standard 25 percent sucrose solution was used in all tests in this series. The comparison stimuli were 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, and 46 percent dextrose.[3]

TABLE II

DEXTROSE CONCENTRATION EQUIVALENT TO 25 PERCENT SUCROSE AND CORRESPONDING SWEETNESS VALUE OF DEXTROSE

Subjects	Mean (conc. dextrose)	SD	SE_M	Sweetness Value (dextrose)
MS	35.3	3.13	.40	71
MC	35.4	2.76	.35	71
MM	34.8	2.40	.31	72
GS	36.2	1.68	.22	69
GF	35.9	3.51	.45	70
All *Ss*	35.5			71

The results show greater variability than in Series 1, and this might be expected. Discrimination in terms of absolute differences is poorer at the higher ranges, and therefore the SD would be expected to increase as it did for all *Ss*. Relative sweetness values for dextrose are higher than in Series 1 for all *Ss*, the mean having risen from 65 to 71. This is the first indication that the relative sweetness of dextrose is not the same at all concentrations.

TABLE III

DEXTROSE CONCENTRATION EQUIVALENT TO 40 PERCENT SUCROSE AND CORRESPONDING SWEETNESS VALUE OF DEXTROSE

Subjects	Mean (conc. dextrose)	SD	SE_M	Sweetness Value (dextrose)
MS	48.6	3.21	.41	82
MC	47.9	3.12	.40	83
MM	47.8	2.85	.37	84
GS	48.2	3.11	.40	83
GF	48.2	3.66	.47	83
All *Ss*	48.1			83

[3] Since percent intervals cannot be assumed to represent equal sense differences at various concentrations, there is no reason why the same interval should intervene between the solutions of one series and another.

Series 3.—A 40 percent sucrose solution was used as a standard. The comparison stimuli were 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, and 55 percent dextrose.

The agreement of the means from one *S* to another is fairly close considering the high concentration of sugars used. In this case the dextrose series ran as high as its saturation point. The 55 percent dextrose solution had a tendency to crystallize slightly, but this fact did not affect the equality judgments, since the solution lay beyond the equality range.

The sweetness value for dextrose again shows a tendency to rise, the average going from 71 to 83. In this case the tendency may be attributed partly to qualitative changes in the dextrose. The sweetness values for dextrose indicate clearly that there is no single value which can adequately express the sweetness value for this sugar. The sweetness of dextrose as a function of the concentration is apparently quite different from that of sucrose.

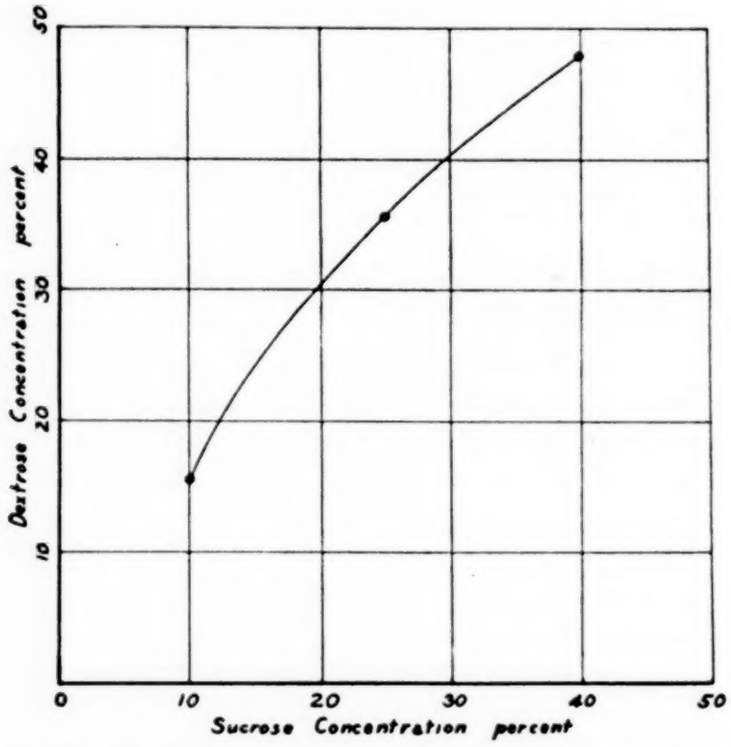

FIG. 1. The concentration of dextrose equivalent in sweetness to sucrose

The SD is again higher in the case of each *S*, indicating poorer sensitivity at the high concentrations used in this series.

Since the distributions of equality judgments in this series are definitely bimodal, SE_M is not a good meausre of reliability. When SE_M was obtained for means of ascending and descending runs taken separately it was found not to exceed .20.

A summary of the results of Experiment 1 is given in Fig. 1.

Experiment 2

A 40 percent sucrose solution was used as a standard. Comparison stimuli were 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, and 54 percent solutions of a two-thirds sucrose, one-third dextrose mixture. Results of this experiment are shown in Table IV.

TABLE IV

SUCROSE-DEXTROSE MIXTURE EQUIVALENT TO 40 PERCENT SUCROSE AND CORRESPONDING SWEETNESS VALUE OF THE MIXTURE

Subjects	Mean (conc. mixture)	SD	SE_M	Sweetness Value (mixture)
MS	42.6	6.32	.81	94
MC	42.7	5.84	.75	94
MM	42.7	5.01	.64	94
GS	42.9	5.31	.68	93
GF	42.6	5.72	.73	94
All *Ss*	42.7			94

One might logically expect that the mixture equivalent to a 40 percent sucrose solution would lie between 40 percent and 48 percent (the latter figure being the dextrose equivalent of a 40 percent sucrose solution as determined by inspection from Fig. 1). Since the mixture is two-thirds sucrose one might expect the concentration of the equivalent mixture to lie nearer the sucrose than the dextrose concentration. By linear interpolation the concentration of the equivalent mixture is found to be 42.68 percent. This agrees very closely with the obtained mean value 42.7. Both the obtained and the calculated equivalent solutions of the mixture have a sweetness value of 94.

SUMMARY

1. A modified method of limits involving successive comparison is found to yield reliable results when applied to the problem of relative sweetness.

2. Sweetness values of dextrose when based upon sucrose as a standard equal to 100 are found to vary with the concentration. At a concentration of sucrose of 10 percent, dextrose has a relative sweetness value of 65. At a concentration of sucrose of 25 percent, the dextrose sweetness value rises to 71 and at a sucrose concentration of 40 percent, dextrose reaches a sweetness value of 83.

3. The relative sweetness of a two-thirds sucrose, one-third dextrose mixture is found to be very close to a value obtained by calculation from the separate sweetness values of the two sugars at a concentration equal to that of the mixture.

4. Qualitative changes in dextrose are apparent at high concentrations. Such changes suggest that sweetness may not be the only taste quality influencing S's judgments.

(Manuscript received October 18, 1947)

References

1. Biester, Alice, Wood, Mildred, & Wahlin, Cecile. Carbohydrate studies. I. The relative sweetness of pure sugars. *Amer. J. Physiol.*, 1925, **73**, 387–396.
2. Carr, C. J., Beck, Frances F., & Krantz, J. C. Sugar alcohols. V. Chemical constitution and sweet taste. *J. Amer. chem. Soc.*, 1936, **58**, 1394–1395.
3. Moncrieff, R. W. *The chemical senses.* New York: Wiley, 1946.
4. Onsdorff, T. Corn sugar utilization in the manufacture of horticultural food products. Unpublished Master's thesis, Mass. State College, 1935.
5. Renner, H. D. Some distinctions in the sweetness of sugar. *Confectionery Production*, 1939, **5**, 255–256.
6. Willaman, J. J. Tests on the relative sweetness of sucrose and levulose. *Z. ver. deut. Zuckerind.*, 1927, 365–367.

THE RATE OF BLINKING DURING PROLONGED VISUAL SEARCH

BY A. CARPENTER

Of the Medical Research Council's Applied Psychology Research Unit in the Psychological Laboratory, Cambridge, England

Introduction

There has recently been some controversy as to whether or not the frequency of blinking can be used as a sensitive index of visual efficiency. Luckiesh and Moss (**6**) have used it, together with other criteria, to investigate particularly the effects of lighting, type of print, etc., during reading, but Luckiesh admits (**7**) "—when the experimental procedure involves not normal reading, but a critical examination of printed matter with the attendant requirements for speed and accuracy, the observed blink-rates usually are inconclusive." Bitterman (**1**), however, concludes that "—present evidence compels us to deny the validity of conclusions based upon this index, either for visual work in general, or for 'normal' reading in particular."

The experiment now to be described, was performed to determine the effect upon the rate of blinking, of visual work of a type quite different from that used by these, and other workers; and it was found that there was a considerable rise in the rate as the test proceeded.

Procedure

The task upon which the *S* was engaged was the Clock Test, devised by Mackworth (**8**) to study the efficiency of visual watchkeepers or look-outs, such as Radar Operators, under conditions where signals were very few, and effects of boredom or fatigue might have been expected. The *S* sat in an electrically screened cubicle, through a window in the front of which he watched a black pointer describing a circle on a plain white background in small steps like the seconds hand of a clock. The pointer was five in. in length, and made 100 steps per complete revolution. Each movement was thus 3.6° in extent, and they occurred regularly at the rate of one per sec. Twelve times, at irregular intervals, in each half hour, this pointer moved through twice the usual distance, i.e., 7.2°, in the same time, and these double-length movements were the 'signals' for which the *S* had to watch. As soon as he noticed each one of these double-length movements he pressed a response key. The display was at a distance of seven feet from the *S*, and illuminated to approximately 10 foot candles by two strip lights, on either side of the window through which the *S* watched. The pointer was painted matt black, and the background was an unmarked sheet of plain white paper. The inside of the cubicle was in semi-darkness, the strip lights illuminating the display being shaded to prevent direct light entering the window. The latter subtended 25° at the *S*'s eye.

The blinks were recorded automatically by a Matthews inkwriting oscilloscope, fed by a five stage battery operated amplifier, of which the first stage was of the push-pull anti-interference type. The time constant of the first resistance-capacity coupling in this amplifier was reduced to about 0.1 sec., so that each blink was recorded as a biphasic movement of the oscillograph pen.

The electrodes consisted of eight-mm. diameter discs of copper foil, of which the central part was depressed to a depth of 0.5 mm. for four mm. diameter. In this central part, a small drop of electrode jelly was placed. The edge was stuck to the *S*'s skin with theatrical spirit gum. The two electrodes were placed centrally just above the eyebrow and on the lower orbital margin. Thus, horizontal eye-movements produced very little output voltage, but the vertical movement, occurring with each blink, produced a small 'square wave' voltage, which, being differentiated by the short time constant coupling in the amplifier, gave the biphasic record mentioned. It is clear that this unidirectional voltage, of about 100 μ-volt, occurring with each blink is due to the small upward movement of the eye, redistributing the corneo-retinal potential in the surrounding tissues, and not to action potentials in the orbicularis oculi (3, 4). In fact, action potentials in the facial muscles gave rise to considerable interference with the record, for example, on frowning or lifting the eyebrows; but the amount of each record lost by this means was very small.

A continuous record of the eyeblinks was thus taken for the whole two hour period of each experimental run. Twenty *S*s each did a two hour run, and for each *S* the mean number of blinks per min. in each half-hour was calculated and is shown in Table I.

TABLE I

MEAN NUMBERS OF BLINKS PER MINUTE FOR EACH *S* IN EACH HALF-HOUR

Subject	Blink Rate per Min.				Mean	Standard Deviation
	Half-Hours					
	1	2	3	4		
1	8.4	16.9	16.2	17.2	14.7	8.73
2	7.3	14.2	15.4	16.6	13.4	6.05
3	10.0	15.6	19.4	19.6	16.2	4.68
4	6.6	10.4	7.6	10.7	8.8	3.45
5	12.3	15.3	15.4	12.5	13.9	5.9
6	30.8	28.3	21.4	20.0	25.1	12.9
7	7.8	9.7	12.3	11.6	10.4	4.66
8	14.9	20.3	21.3	19.7	19.1	7.49
9	28.8	30.6	33.1	36.5	32.2	9.75
10	7.9	10.0	16.7	16.9	12.9	5.54
11	17.8	20.7	21.9	25.6	21.5	7.98
12	8.3	15.4	16.9	12.6	13.3	5.84
13	31.9	32.6	35.0	39.2	34.7	7.93
14	12.1	12.3	12.5	13.7	12.7	4.6
15	21.4	25.1	23.0	24.1	23.4	4.1
16	17.1	32.3	26.5	36.4	28.1	6.28
17	21.6	29.4	29.9	30.1	27.8	9.82
18	12.2	16.9	18.2	19.7	16.8	4.36
19	59.1	60.5	42.2	83.7	61.4	14.4
20	11.1	21.9	16.9	30.3	20.1	9.5
Mean	17.4	21.9	21.1	24.8	21.3	

RESULTS

The *mean rate* of blinking is seen to increase especially between the first and second half hours, with a further slight increase in the last half-hour. This rise is statistically significant, its variance ratio taken from the analysis of variance given in Table II being 7.06. Further, of the 20 *S*s, only one, No. 6, shows a fall in rate, and one other, whose rate was very high, No. 19, was erratic. The other 18

TABLE II

ANALYSIS OF VARIANCE OF THE HALF-HOUR MEANS GIVEN IN TABLE I

Source of Variation	Degrees of Freedom	Sum of Squares	Variance	Variance Ratio	1% Limit
Between half-hours	3	512	170.7	7.06	4.16
Between subjects	19	10,908	574.1	23.75	2.23
Residue ≡ error	57	1,378	24.17		
Total	79	12,798			

all show a clear rise. This upward trend is clearly shown in Fig. 1 in which the ratios for the second, third and fourth half-hours are shown as percentages of the first half-hour rate, for each *S*. The clock scores in this experiment were unfortunately invalidated by the presence of a supplementary auditory cue. But the results usually obtained, as the number of signals missed in each half-hour, show an upward trend very similar to that of the blink rate, found here (Mackworth, **8**).

The *variability* in the rate of blinking is, of course, considerable, between different *S*s. But, also, the variability of any one *S*'s rate,

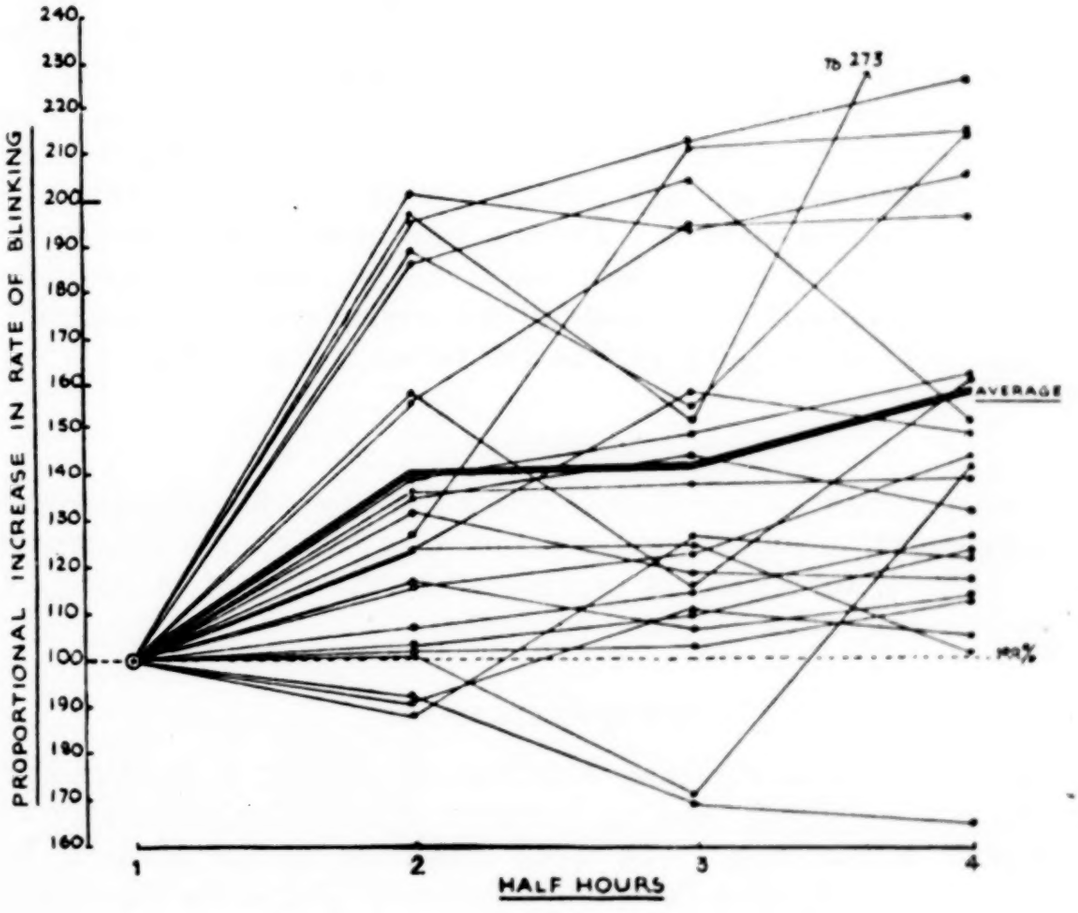

FIG. 1. Rates of blinking during clock test, initial rate being taken as 100

from minute to minute, is very marked. This is also shown in Table I, where the last column gives the standard deviation of the numbers of blinks per min., calculated for each *S* over the whole two-hour test.

Discussion

This experiment clearly supports Luckiesh's contention that, at least under some circumstances, the rate of blinking can be used as a gauge of eye-strain or fatigue. This task is, of course, very different from the reading tasks used by him, or by Hoffman (**5**) and would probably be expected to cause a greater change in the rate, than would reading. It is known that any person's blink rate depends very much upon what he is doing, being greater during conversation, for example, than while reading. The suggestion that this present task is more severe than reading is supported by the relatively very low blink rates found, averaging 21.3 per min. for all the *S*s, compared with, for example, 35 per min. recorded by Luckiesh and Moss, for *S*s engaged on reading. Similarly the overall percentage increase in rate found in the experiment, 43 percent, is somewhat greater than these workers obtained, namely 31 percent.

The results of some previous workers (**9**) have been criticised on the ground of statistical unreliability. The present results are very reliable, having a variance ratio of 7.06, giving $P < .01$. But it is clear that the variability of the blink rate, not only from *S* to *S*, and from one occupation to another, but from minute to minute in any one *S* engaged in a set task, makes it necessary to be cautious in basing conclusions upon it. Thus, a rate based upon counting the blinks for a period of five min. only will be liable to considerable error, and it is possible that some of the present confusion concerning the reliability of the blink rate as a criterion is due to this.

Summary

An experiment is described, in which the rate of blinking increased by 43 percent, during the performance of a two-hour watchkeeping task. This is submitted as evidence that the rate of blinking can be used as a criterion of visual efficiency.

Acknowledgments

I owe thanks to Professor Bartlett and Dr. K. J. W. Craik for advice and assistance, and to Professor Adrian for the use of the Matthews oscillograph. The work was done in the Medical Research Council's Applied Psychology Research Unit in the Psychological Laboratory, Cambridge, England.

References

1. Bitterman, M. E. Frequency of blinking in visual work, a reply to Dr. Luckiesh. *J. exp. Psychol.*, 1947, **37**, 269–270.
2. Bitterman, M. E., & Soloway, E. Frequency of blinking as a measure of visual efficiency. Some methodological considerations. *Amer. J. Psychol.*, 1946, **59**, 676–681.
3. Bitterman, M. E. Electromyographic recording of eyelid movements. *Amer. J. Psychol.*, 1945, **58**, 112–113.
4. Hoffman, A. C., Welman, B., & Carmichael, L. A quantitative comparison of the electrical and photographic technique of eye-movement recording. *J. exp. Psychol.*, 1939, **24**, 40 sq.
5. Hoffman, A. C. Eye-movements during prolonged reading. *J. exp. Psychol.*, 1946, **36**, 95–118.
6. Luckiesh, M., & Moss, F. K. *The science of seeing.* N.Y.: D. van Nostrand Co., 1937.
7. Luckiesh, M. Reading and the rate of blinking. *J. exp. Psychol.*, 1947, **37**, 266–268.
8. Mackworth, N. H. The breakdown of vigilance during prolonged visual search. *Quart. J. exp. Psychol.*, 1948, **1**. (In preparation)
9. Tinker, M. A. Reliability of blinking frequency employed as a measure of readability. *J. exp. Psychol.*, 1945, **35**, 418–424.

APPARENT MOVEMENT IN RELATION TO HOMONYMOUS AND HETERONYMOUS STIMULATION OF THE CEREBRAL HEMISPHERES [1]

BY J. A. GENGERELLI

University of California, Los Angeles

Introduction

Apparent visual movement arises when two discrete stimulus points, p_1 and p_2, are presented in temporal succession under conditions of optimum spatial and temporal separation. If p_1 is given first and p_2 second, the movement is from p_1 to p_2. The position of p_2 in space with respect to p_1 determines the direction of the movement from its point of origin, for, if the position of p_2 were rotated 90 degrees, the movement vector would suffer a similar rotation. Hence, the phenomenon of apparent movement constitutes one of the most compelling examples of gestalten, in that the direction of the movement cannot be established until the second of the two points, p_2, has been presented. Thus, the phenomenon in its entirety is dependent upon p_1 and p_2 and is modified by any suitable change in either or both: that is, the excitations created by p_1 and p_2 form a *system* and hence are interrelated.

In the present investigation, a preliminary attempt is made to study the degree of interrelationship between cerebral excitations as influenced by their locus in the cerebral mass. In particular, is the interrelation stronger if p_1 and p_2 give rise to excitations in the same cerebral hemisphere than if they cause excitations in disparate hemispheres? Let it be supposed that at time t_1 we present the two visual points $_{\bullet p_1}$ $^{\bullet p_1'}$ and at time t_2 we present the two visual points $^{\bullet p_2}$ $_{\bullet p_1'}$. These four points are so related one to another gemetrically as to form the four corners of a square: $\bullet p_2$ $\bullet p_1'$. It is clear that the $\bullet p_1$ $\bullet p_2'$ spatial and temporal parameters involved in the presentation of these stimuli are such as to make equally probable (a) movement from p_1' to p_2' and from p_1 to p_2, and (b) movement from p_1 to p_2' and from p_1' to p_2. In the first case we should have reciprocal *vertical* movement; in the second case, reciprocal *horizontal* movement.

[1] Reported in part at the Philadelphia meeting of the American Psychological Association, 1946.

Let it be supposed that the visual fixation point is at the middle of the square constituted by the four points and that the dimensions of the square are such that points p_1' and p_2' fall on one side of the retinal midline, and points p_1 and p_2 fall on the other. Then the stimulation from the primed points will go to one cerebral hemisphere and that from the unprimed points will go to the other.[2] If now, in this situation, we obtain predominantly reciprocal *vertical* apparent movement, we are forced to conclude that the interrelation between the homonymous points is stronger than between the heteronymous points; if we obtain predominantly reciprocal *horizontal* apparent movement, our conclusion is reversed; finally, if no dominance of one over the other is indicated, we may conclude that the two types of interrelation have approximately equal strength.

As a check, we may envisage an experiment in which the visual fixation point is to one side of the square of four points. In this case all four points would be homonymous.

Apparatus and Method

The apparatus consisted of two rectangular tunnels 3.5 in. by 5 in. placed at right angles to each other in such wise as to form a short-stemmed T. One objective was placed at the far end of the cross-piece and the other objective at the far end of the stem. The *S*'s eye was at the near end of the cross-piece. At the point of juncture of the two tunnels, a half-mirror was so placed that light from the objective at the far end of the stem was reflected from the mirror, whereas light from the objective at the far end of the cross-piece was transmitted through the mirror, to the *S*'s eye. The distance of the two objectives from the observation point was identical: 12 in.

The objectives were in the nature of white cards three in. by three in.; each bore three black dots, 0.5 mm. in diameter, arranged along an imaginary diagonal line such that the two end dots formed the diagonal corners of a five mm. square. When the two cards were illuminated simultaneously, the *S* saw four dots arranged in the form of a square; in the center of the square the *S* saw a single dot which served as a fixation point.

Light was reflected from the surface of the cards, and was provided by 10-watt incandescent bulbs. Two such bulbs were used as a source of illumination for each card. The bulbs were enclosed in an especially constructed container which projected the light through an opening in the tunnel near the objective. The set-up was so arranged that only light reflected from the surface of the cards reached the *S*'s eye.[3] The cards were illuminated in an alternating manner by means of a circuit provided with two rotary switches, in such wise that the interval of darkness in the visual field equalled the interval of illumination; the value of this interval was about .10 sec. It was found that this interval gave the most palpable and satisfactory apparent movement.

The *S* viewed the objective monocularly through a circular opening about .75 cm. in diameter in the near end of the cross-piece of the tunnel. He was provided with a crude chin rest, and his eye was about one inch from the opening. He saw under conditions of illumination a circular field.

Each *S* was given a brief description of the stimulus cards and told that two pairs of diagonal dots would be alternately flashed on and off; that under these conditions it was usual for the *S* to see the dots move from one position to the other. The positions of the dots were illustrated

[2] It is assumed that the fovea suffers bisection as does the periphery of the retina.

[3] Considerable effort was made to obtain commercial 'instantaneous' bulbs with very little lag in illumination and extinction. A three-watt gas bulb with large electrodes was procured which fulfilled these requirements and was used in the preliminary experiments. It was found, however, that these bulbs provided too little illumination. It was finally decided to use a 10-watt incandescent bulb for the purpose. Since the time values involved in the experiment were large, the lag in illumination and extinction was found to present no problem.

to him on paper, and the paths of possible horizontal and vertical motion indicated. It was pointed out also that possibly for him the movement might be neither horizontal nor vertical, but rotary or of some other form. It was found necessary to 'indoctrinate' the *Ss* in this manner, in order to establish a 'priority of observation.' When this was not done, it was found that the *Ss*' experiences were phenomenologically less structured and that it was much more difficult to obtain quantifiable data.

The *Ss* were instructed: (1) to keep the fixation point in the center of the field at all times; it was pointed out to them that any change in the position of the point from dead center indicated movements of the head which were not desirable; (2) while fixating the point in the center of the field at all times, to report any observations of movement of the other points under the following heads, (a) vertical movement and (b) non-vertical movement.[4]

The *S* was given a compact hand switch which actuated a cumulative electric clock. He was instructed that he was to press the switch when the movement was vertical and to release it when the observed movement was non-vertical. Each period of observation was two min. A five-min. period of practice in observation and manipulating the switch was given prior to the experiment proper, and five-min. rest intervals were provided between periods. All experimentation was carried out in a dark room. After each observation period, protocols were obtained; in particular, the *S* was questioned regarding what he saw during the times he reported non-vertical movement.

Experiment I

In this experiment, eight *Ss* were used; all had had extensive psychological training. A series of four observations were made, which will be referred to as A, B, C, and D. In A, the fixation point was in the center of the 'square'; in B, it was three mm. to the right; in C, it was three mm. below; and in D, it was four mm. diagonally below the lower right dot of the 'square.' These observations were not made in the same order by the *Ss*; rather, the sequences were 'randomized' for each *S* in order to eliminate any observational bias.

The results are shown in Table I, in terms of the amount of time during which vertical movement was seen and the amount during which nonvertical movement was seen. Examination of the protocols showed that probably 95 percent of the non-vertical movement reported by the *Ss* was in the nature of horizontal reciprocating motion.

A calculation was made of the reliability of the differences for each of the four series of observations on the basis of the differences between the vertical and the non-vertical scores for each *S*. The results are as follows:

For A: t = 4.12; reliable at the one percent level
For B: t = .685; not reliable at the five percent level
For C: t = 10.795; reliable at the one percent level
For D: t = .76; not reliable at the five percent level

From these results it is evident that vertical movement predominated over horizontal movement in A and C; that is to say, in

[4] The *Ss* were given a careful description as to what was meant by *vertical* and *non-vertical*. The former was described as movement of the points from ceiling to floor; the latter as any other kind of movement, such as horizontal (from one side of the room to the other), diagonal, circular, etc.

those observations where homonymous cerebral excitations were pitted against heteronymous cerebral excitations, the inter-relations between the former were definitely stronger. Series B and D served as controls in this experiment, in that in these two series all four dots presumably caused excitations in the same cerebral hemisphere.

TABLE I

TIME IN SEC. DURING WHICH VERTICAL (V) AND NON-VERTICAL (N. V.) MOVEMENTS WERE SEEN BY EACH *S* UNDER EACH CONDITION, EXPERIMENT I

	A		B		C		D	
	V	N. V.	V	N. V.	V	N. V.	V	N. V.
1	73	47	0	120	89	31	29	91
2	91	29	65	55	120	0	107	13
3	120	0	120	0	120	0	120	0
4	59	61	0	120	120	0	0	120
5	110	10	0	120	105	15	109	11
6	73	47	120	0	113	7	111	9
7	117	3	63	57	92	28	2	118
8	120	0	10	110	120	0	120	0
Mean	95.4	24.6	47.3	72.7	109.9	10.1	74.8	45.2

Table I is worthy of study by virtue of the fact that at least two *S*s, namely Nos. 3 and 6, reported predominance of vertical movement regardless of the position of the fixation point. Whether this is attributable to the possibility that they ignored directions and fixed the center of the 'square,' or rather represents a factor of systemmatic individual difference, it is impossible to say. There are in the table other notable individual differences of a cognate character.

At the end of the series of observations each *S* was asked once more to report under condition A. In this instance, however, the *S*, after having watched for some seconds in the normal manner, was asked to bend his head to one side some 90 degrees and repeat his observations. In all instances the *S*s reported an analogous rotation of the direction of apparent movement. If under the normal condition, the movement had been from ceiling to floor, it now became a movement from wall to wall; and *vice versa*.

EXPERIMENT II

All the essential conditions of the preceding experiment obtained here. For this experiment however, 14 new *S*s were used, all students in an undergraduate course in psychology.

The stimulus cards were changed in the following manner. Small colored discs, three m m. in diameter, were used instead of black dots. The distance between the stimulus points was increased, so that the four now formed a 'square' 10 mm. on the side. Red, green, yell ow, and blue colors were used; in Condition A, the upper left, upper right, lower left, and lower right dot were, respectively, Y, R, B, and G, while in Condition B they were G, R, G, and R. A smal

black dot was still used as a fixation point, however, and at the same distance from the 'moving' points as in the corresponding conditions in the preceding experiment.

A change was made also in the fixation device, to make the task of visual fixation more tangible for the observer. A small black dot was placed on the half-mirror. The *S* was now instructed to keep this dot and the fixation point in alignment at all times, in such wise that but one black dot was visible in the visual field. Any slight movement of the head would produce two black points in the field, thus providing him with a cue for the necessary adjustment.

In this experiment observations were carried out only under conditions A and B.

Results are given in Table II. Examination of the protocols again shows that the non-vertical movement reported by the observers was predominantly horizontal in nature, although not to the same overwhelming degree as in Experiment I. Some reported circular movement, others movement along a diagonal. Still others reported occasionally a break-down of the movement pattern.

TABLE II

TIME IN SEC. DURING WHICH VERTICAL (V) AND NON-VERTICAL (N. V.) MOVEMENTS WERE SEEN BY EACH *S* IN CONDITIONS A AND B, EXPERIMENT II

	A		B	
	V	N. V.	V	N. V.
1	93	27	100	20
2	75	45	72	48
3	110	10	88	32
4	93	27	98	22
5	116	4	96	24
6	27	93	39	81
7	120	0	120	0
8	120	0	44	76
9	48	72	0	120
10	116	4	0	120
11	100	20	0	120
12	80	40	81	39
13	60	60	97	23
14	46	74	15	105
Mean	86	34	60.71	59.29

The results are recorded in seconds.
V = vertical movement; N.V. = non-vertical movement.

Calculation for condition A gives a $t = 3.15$ which is reliable at the one percent level. No attempt was made to calculate the reliability for condition B, since it is evident from inspection that the difference in time is negligible.

Inspection of Table II again reveals the fact that certain observers reported a great predominance of one kind of movement regardless of the condition under which the observations were made.

As in Experiment I, at the end of the series of observations, the *S*s were asked to report once more under condition A, with the head

inclined 90 degrees, after having watched the movement from a normal position for a short time. All *Ss* except one reported a corresponding rotation of movement. This change in the direction of the motion uniformly impressed these psychologically naive *Ss*, in particular by virtue of the fact that the discs were of different color, and the change was therefore the more spectacular. Thus, if in the immediately preceding normal position the *S* had seen vertical movement, this movement would be between the red and green discs, and the yellow and blue discs, respectively. Upon rotating the head, the movement took place between the red and yellow discs, and the blue and green discs, respectively.

In this experiment, as in the preceding one, when the *S* reported rotation of movement, he was asked to make a definite attempt to restore the movement to its original direction. All reported a remarkable stability of the movement and found themselves unable to change the direction at will.

The one *S* who formed the exception reported such frequent shifts of movement from vertical to non-vertical when in a normal position that the test could not be made for him.

Miscellaneous Observations

(a) In the course of these experiments it was noted that the rapidity of fluctuation between vertical and non-vertical movement varied considerably from *S* to *S*, although no attempt was made to measure this aspect of the phenomenon. The impression was gained, however, that under conditions when all four stimulus points involved homonymous cerebral excitations the frequency of shift was much greater than when homonymous excitations were pitted against heteronymous excitations.

(b) In Experiment II a number of *Ss* complained spontaneously that they "couldn't tell just exactly when, for example, the red dot changed into the green dot, and *vice versa*."

(c) The writer spent, altogether, a number of hours acting as *S* in this experiment. At the beginning, for any given condition of observation, he experienced the shifts in the direction of movement reported by the other *Ss*; as time went on, however, condition A gave uniformly vertical motion, without shift, and condition B gave similarly uniformly horizontal motion. Many times, while observing under A and seeing consistent vertical motion, the writer deliberately turned his fixation to the right of the field of the moving points, thus simulating condition B. Almost at once the movement would switch from vertical to horizontal. Upon restoring the line of regard, the motion would snap back into the vertical type.

(d) Another interesting fact reported by some *S*s (noted also by the writer) was that a given form of motion, e.g., the vertical, was often seen in two ways, *viz.*: (a) as a simple straight up-and-down reciprocal motion of the dots or, (b) as the vertical up-and-down motion of the end-points of an imaginary diagonal rod rotating back and forth on a pivot point at the center. This was observed particularly in condition A, where there was an actual dot in the center of the movement field.

Discussion

It was deemed unnecessary to perform a parallel experiment in which the *S*s were asked to report their observations in terms of 'horizontal' and 'non-horizontal' movement. While it may be conceded that the form of the instructions given might affect the *S*'s reports, this possible influence is irrelevant to the purpose of the experiment. The experiment was designed to determine whether any change in the direction of seen movement accompanies a change in the fixation point; in particular, to observe the nature of the seen movement (a) under conditions when homonymous stimulation only was present, and (b) when both homonymous and heteronymous stimulations were present. Any statistically reliable differences between the results given by these two conditions can be attributed only to the presumed change in cerebral excitation, and hence our finding that homonymous excitations are more closely interrelated than heteronymous ones cannot be impugned on these grounds.

Furthermore, the results obtained at the end of each experimental session, when the *S*s were asked to observe once more under condition A and then to bend the head to one side approximately 90 degrees, serve as a useful check. In all instances, the *S*s reported a corresponding rotation in the direction of the seen movement. In this connection, it is necessary to bear in mind that the definition of vertical movement given to the *S* was that of *movement from ceiling to floor*, and not as movement from forehead to chin.

Summary

Two pairs of visual stimulus points were presented to *S*s under conditions producing the phenomenon of apparent movement, but in such wise that two radically different forms of seen movement were available, *viz.*, vertical and horizontal.

By altering the position of the visual fixation point relative to the points in question, an attempt was made to determine whether homonymous cerebral excitations were more intimately interrelated than heteronymous cerebral excitations. In the event this were true,

one form of the movement should predominate over the other under certain defined conditions.

In the case where both homonymous and heteronymous forms of cerebral excitation were simultaneously present it was found that that form of the seen movement was predominant which had its base in the homonymous cerebral excitations. On the other hand, when only homonymous cerebral excitations were present, there was no clear predominance of one type of movement over the other.

It is concluded that two excitations existing in the same cerebral hemisphere are more strongly interrelated than two excitations existing in different cerebral hemispheres.

(Manuscript received September 3, 1947)

INFANT SPEECH: SPEECH SOUND DEVELOPMENT OF SIBLING AND ONLY INFANTS

BY ORVIS C. IRWIN

Iowa Child Welfare Research Station

In an article (1) previously published it was reported that the speech of infants living in an orphanage during the first six months of life is retarded in comparison with infants living in their own homes. In another report (2) it was shown that low grade feebleminded children with an average age of five years are comparable in speech sound status to year and a half old babies. Another possible variable in the speech development of infants is the presence in the family of older siblings.

The purpose of the present article is to present the central tendencies of speech sould development during the first two and a half years of life for only infants and for infants with older siblings. The

TABLE I

PHONEME TYPE MEANS AND FREQUENCY MEANS FOR ONLY INFANTS AND INFANTS WITH SIBLINGS

Bimonthly Age Group	Number of Subjects		Type Means for		Frequency Means for	
	Only	With Sibs	Only	With Sibs	Only	With Sibs
1–2	32	30	6.8	7.5	61.7	58.8
3–4	44	36	11.2	11.0	69.0	64.8
5–6	44	31	12.3	12.2	75.8	71.1
7–8	38	26	14.6	15.9	77.5	74.8
9–10	35	27	15.9	16.0	84.8	76.2
11–12	34	28	19.1	18.1	93.0	89.6
13–14	32	25	19.6	17.8	90.8	84.1
15–16	31	24	20.6	20.3	90.1	95.3
17–18	30	20	21.3	20.6	98.0	97.4
19–20	22	19	22.4	22.7	96.7	99.7
21–22	21	16	22.9	23.2	110.4	102.8
23–24	17	14	25.2	23.4	119.5	109.9
25–26	19	13	26.3	25.2	141.4	132.5
27–28	12	12	24.8	27.4	130.0	150.8
29–30	10	9	28.1	26.1	150.8	161.3

data are analyzed in terms of phoneme type and phoneme frequency and are included in Table I. The table is read as follows: For the first bimonthly age group the number of only infants is 32 and of infants with siblings it is 30. The type mean for only infants is 6.8 and for infants with siblings it is 7.5. The frequency mean for only infants is 61.7 while for infants with siblings it is 58.8.

Equations have been derived from the data. Theoretical curves

have been plotted from the equations and are shown in Figs. 1 and 2. The curves for phoneme type are parabolic arcs. The equation for these curves for only infants is $P_T = 8.0\ A^{0.451}$, and for infants with sibs it is $P_T = 7.9\ A^{0.446}$. Thus the constants in the phoneme type equations are practically identical. The curves for phoneme fre-

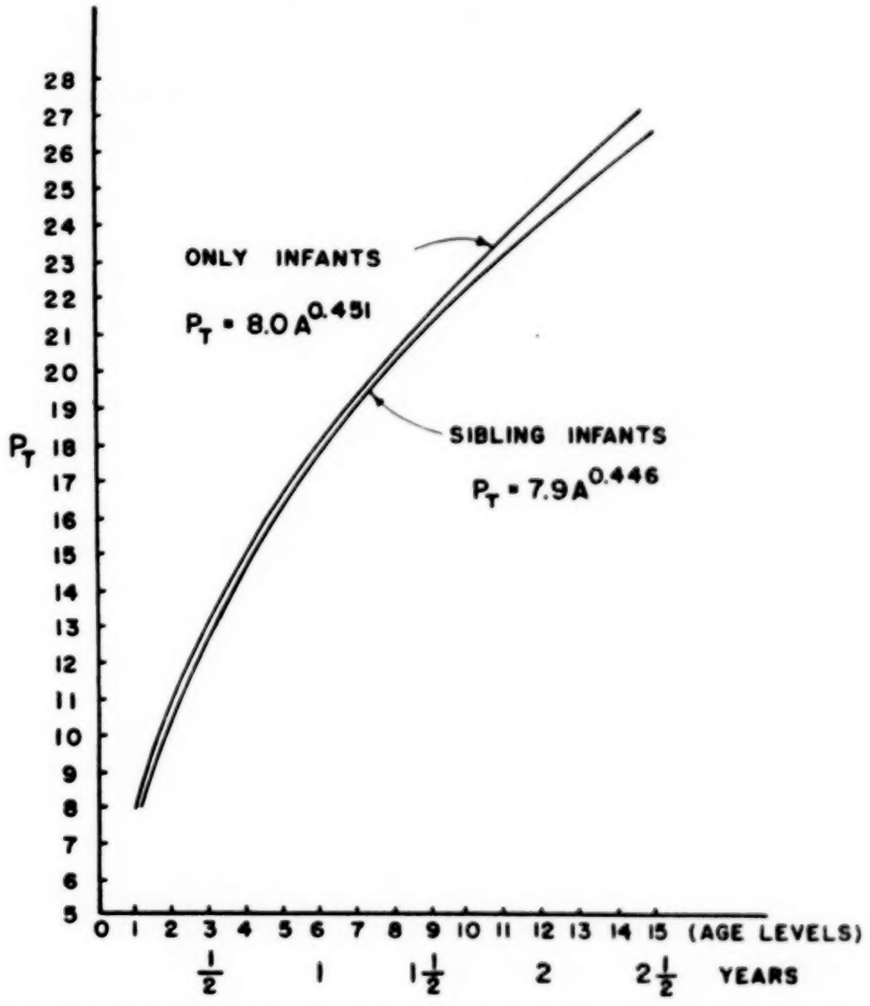

FIG. 1. Curves for phoneme type fitted to data for sibling and only infants

quencies are exponentials. The equations for the frequency curves are $P_F = 61\ e^{.055A}$ for the only infants, and $P_F = 57\ e^{.063A}$ for the infants with siblings. The constants in the equations for phoneme frequency also are practically the same. The pairs of curves show these results graphically.

It is thus a reasonable inference as far as measures of the central tendencies of these two criteria are concerned that the presence of siblings in the family has negligible effect upon the speech sound development of infants.

SUMMARY

A possible variable in the speech sound development of infants is the presence of older siblings in the family. This article is an anal-

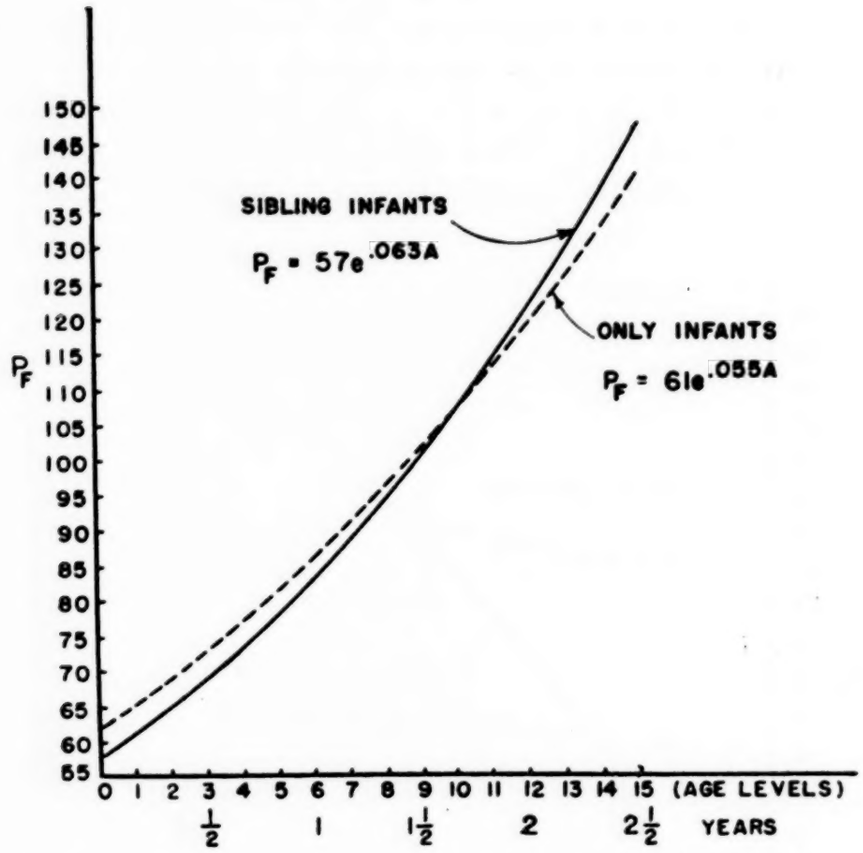

FIG. 2. Curves for phoneme frequency fitted to data for sibling and only infants

ysis of the means for only infants and for infants with siblings for the first two and a half years of life. The data are analyzed in terms of phoneme type and phoneme frequency.

Equations derived from the means of phoneme type are $P_T = 8.0\,A^{0.451}$ for only infants, and $P_T = 7.9\,A^{0.446}$ for infants with siblings. The equations derived from the frequency values are $P_F = 61\,e^{.055A}$ for only infants and $P_F = 57\,e^{.063A}$ for infants with siblings. It is seen that constants in the equations for the type data are practically the same. The constants in the frequency equations likewise are fairly close. Thus, an examination of the data in terms of central tendencies in development reveals little difference between the two graphs of infants. Theoretical curves are included in the article.

(Manuscript received October 11, 1947)

REFERENCES

1. BRODBECK, A. J., & IRWIN, O. C. The speech behavior of infants without families. *Child Develpm.*, 1946, 17, 145–156.
2. IRWIN, O. C. The developmental status of speech sounds of ten feebleminded children. *Child Develpm.*, 1942, 13, 29–39.

STUDIES IN SPACE ORIENTATION. III. PERCEPTION OF THE UPRIGHT IN THE ABSENCE OF A VISUAL FIELD

BY H. A. WITKIN

Brooklyn College

AND

S. E. ASCH

Swarthmore College *

In earlier studies (1, 2) it was established that the main directions of the visual field play a decisive role in the perception of the vertical and horizontal. The aim of the experiments described here was to determine how the upright is established in the absence of a surrounding visual field. Since exclusion of the field leaves only the body as a basis for judgment, the problem was to determine how effective orientation can be when based upon postural factors.

To this end a series of experiments was carried out in which the visual framework was eliminated by the simple expedient of working in a completely darkened room. The sole visual datum before the *S* was a luminous rod, which he had to try to adjust to the true vertical and horizontal. In order to evaluate the specific role of postural factors, the positions of the head and body were varied systematically, and the accuracy of judgments at each position determined.

I. Procedure

A. Experiment I

In this experiment the *head alone* was tilted.

To insure the same angle of tilt with each *S*, and to maintain this angle throughout a test, the following simple technique was employed. The *S* stood erect in a tall, rectangular box, 74 in. high by 22 in. wide by 12 in. deep, open at front and top. A right triangle, which could be moved up and down to accommodate the *S*'s height, was attached to the left or right side of the box, as required. This triangle was lowered onto the *S*'s shoulder, and he rested his head against the plane formed by its hypotenuse. In this way it was possible to maintain the head at about a 45° angle.

The rod used in establishing the V (vertical) and H (horizontal) was 39 in. long and one in. wide. The surface of the rod facing the *S* was coated with luminous paint, which was brightened up before each experiment by passing a light over it. The rod was pivoted at its center; and attached to the other end of the shaft on which it rotated was a protractor. Since this protractor turned with the rod against a stationary pointer, the amount of movement of the rod could be read off directly in degrees. This unit stood about five feet in front of the *S*, with the pivot point of the rod four feet off the ground and in line with the median plane of the *S*'s body.

* These experiments were performed in the psychological laboratories of Brooklyn College.

Sixteen judgments were obtained from each *S*, eight with the head tilted right and eight with the head tilted left. Of each set of eight judgments, four were for the H and four for the V. When the rod was to be made H, it was initially placed in a V position, and vice versa. Movement of the rod to the required position was carried out by the *E* at the *S*'s direction. On half the trials the rod was moved clockwise, on the other half counter-clockwise. Movement was executed in small steps, and after each turn the *S* was asked to say 'Continue,' until he was satisfied that it was in the correct position.[1] The eight judgments with head tilted right were obtained first, and then the eight judgments with head tilted left. In each part, two H judgments, one with clockwise movement of the rod and the other with counterclockwise movement, were alternated with two V judgments, again one with clockwise and the other with counterclockwise movement.

The *S* was brought into the laboratory blindfolded, and at no time was he permitted to see anything but the luminous rod. When the lights were turned on to take protractor readings, the *S* was required to close his eyes. All this was done to prevent him from locating the axes of the room, which he might then possibly have used in making judgments, even when the lights were off.

Twenty-three *S*s were tested in this situation.

B. *Experiment II*

In this experiment the *S*'s whole body was tilted.

1. *Experiment IIa.*—Body tilts of *42°* left and right were employed. The apparatus used to maintain the body at the given tilt consisted of two long boards attached to each other in the form of an inverted V, the angle between them being such that each board was inclined 42° from the vertical. A baseboard attached to each of these inclined planes, at right angles to it, provided a footrest. To obtain a left tilt or right tilt of the body, it was simply necessary for the *S* to lean against the left plane or the right plane, standing on the footrest attached to it. Lying against the board, the *S* placed his arm around it, so that his body was brought close to the board and his head maintained in a fairly natural relation to the body. Regardless of the side to which the *S* was leaning, his head was in line with the pivot point of the luminous rod, which as before was about five feet in front of him.

For the first part of the test, the body was displaced to the *left*. Using the procedure described for Experiment I, four H and four V judgments were obtained. The body was then tilted to the *right*, and the series repeated.

Thirty *S*s were tested in this situation.

2. *Experiment IIb.*—In this experiment the body was tilted only *28°* from the vertical. A somewhat different procedure was employed than in the two previous experiments. In the first place, the *S* was tested with a left tilt of the body only. In the second place, on each trial the rod was initially displaced 28°, rather than 90°, from the true V or H.[2] In the third place, 12 trials were given instead of eight. On the first six of these, the rod was always moved clockwise from its initial position; on the last six, counterclockwise. V and H judgments were obtained in an alternating sequence. In all other respects, the procedure was the same as in Experiment IIa.

Forty-nine *S*s were tested under these conditions.

C. *Experiment III*

In this situation the body was in a *horizontal* position. The *S* was tested first while lying on his left side, and then while lying on his right side. In each position the head was directly in front

[1] In a preliminary experiment it was found that each *S* tended to have a range of positions in which the rod appeared H (or V). If continuous movement had been used instead of this step-wise procedure, some *S*s might have stopped at the periphery of the range, others overshot into the middle of the range, and so on. This would have resulted in a technique which in itself leads to increased variability in performance.

[2] It will be noted that in Experiments I, IIa, and III the rod was initially placed 90° away from the correct position, whereas in IIb and IV it was placed only 28° away. These latter two experiments were done some time after the others, when it had been recognized that the 90° displacement was excessive. The use of smaller initial displacements was particularly justified in these two experiments where it was used because, with the body tilted only 28°, or else erect, the errors in adjustment of the rod tended to be considerably less than 28°.

of the pivot point of the luminous rod. In order to keep the head aligned with the rest of the body, the *S* extended the arm on the side to which he was tilted, and rested his head upon it. The procedure followed in all details that employed in Experiment IIa.

This test was given to 24 of the *Ss* who had served in that experiment.

D. Experiment IV

Here the *S* was tested while *standing erect.* This situation provided a control for the previous experiments. As in Experiment IIb, the rod was initially displaced 28°. Only the clockwise series of trials was given, providing a total of three H judgments and three V judgments.

Thirty-nine *Ss*, all taken from the group employed in Experiment IIb, were tested.

II. Results

The main results are presented in Table I. Each value represents the mean error for the given condition, or the average amount by which the adjustments of the rod deviated from the true V and H, without regard for the direction of the deviations.

TABLE I

Mean Errors in Establishing the V and H for Each of the Body Positions Employed

Throughout, V and H scores have been combined,* and no account has been taken of the direction of the errors. In Experiment IIb, where the body was displaced 28°, the body-right condition was not used.

Body Position	Mean Errors		
	Head or Body Tilted		Total
	Left	Right	
A. Head or body tilted			
1. Head alone tilted (Expt. I)	6.0°	5.7°	5.9°
2. Body tilted 28° (Expt. IIb)	7.6°	—	—
3. Body tilted 42° (Expt. IIa)	5.5°	6.4°	5.9°
4. Body horizontal (Expt. III)	18.5°	13.7°	16.1°
B. Body erect (Expt. IV)			1.5°

* This was done since results for the two types of judgment were very similar. The mean errors for H and V judgments respectively were as follows: (1) Head alone tilted, 5.9° and 6.0°, (2) Body tilted 28°, 8.5° and 6.7°, (3) Body tilted 42°, 5.7° and 6.2°, (4) Body horizontal, 16.3° and 15.9°, (5) Body erect, 1.4° and 1.5°.

These results show first of all that with the body erect (Experiment IV) the V and H can be established very accurately, although no visual field is present. The mean error for the group as a whole was only 1.5°. Twenty of the 39 *Ss* had a mean error of 1° or less, 33 of 2° or less. The poorest *S* had a mean error of only 4°. In contrast, tilting of the head or body leads at once to errors in judgment of the V and H. With only the head tilted (Experiment I), the mean error was 5.9°. With an extreme displacement of the body

(body horizontal, Experiment III), the mean error reached the very high value of 16.1°. In some individual cases, the errors produced by tilting the head or body were considerably larger than the average figures. The poorest mean score for each body position was as follows: head alone tilted, 12.3°; body tilted 28°, 20.8°; body tilted 42°, 14.3°; body horizontal, 37.6°. It is clear from these results that when the body is upright postural factors are fully adequate for judging the V and H, but that when the body, or even the head alone, is tilted, they provide a less effective basis for judgment.

Tilting of the head or body also produces considerable variability in the judgments made by the *S* under a given condition. This becomes evident from an examination of the individual ranges.[3] The following mean ranges for H judgments, obtained under conditions of *left* tilt of the head or body, are representative. These are summarized in Table II. (a) With head alone tilted, the mean range was

TABLE II

MEAN RANGES FOR H JUDGMENTS

Body Position	Mean Range
A. Head or body tilted	
Head tilted left (Expt. I)	6.0°
Body tilted 28° left (Expt. IIb)	9.6°
Body tilted 42° left (Expt. IIa)	9.9°
Body horizontal on left side (Expt. III)	12.8°
B. Body erect (Expt. IV)	2.4°

6.0°. Of the 23 *S*s tested, five had ranges of more than 10°, and one as high as 17°. (b) With a 28° body tilt, the mean range was 9.6°. Of 49 *S*s, 21 had ranges of more than 10°, and 10 had ranges of more than 15°. The range of one *S* was 25°. (c) With a 42° body tilt, the mean range was 9.9°. Of 30 *S*s, 11 had ranges of more than 10°, and five of more than 15°. The largest range was 31°. (d) Finally, with body horizontal, the mean range was 12.8°. Of 24 *S*s, 15 had ranges of more than 10°, and seven more than 15°. The largest range was 25°. It will be noted that there is a progressive increase in the mean range for these four positions, indicating greater variability with more extreme tilts of the body. In contrast to the variability which is manifested when the head or body is tilted, successive judgments prove very consistent when the body is upright; thus the mean range for H judgments with body erect was only 2.4° which is considerably less than any of the values obtained with head or body

[3] In determining the range for each *S* the difference between his two most extreme scores was taken. Account was taken of the direction of the errors. For example, if the judgment was 12° off the upright in a clockwise direction, at one extreme, and 12° in a counterclockwise direction, at the other extreme, the range was counted as 24°. To obtain the mean range for a given condition an average was taken of the ranges of the individuals tested in that condition.

tilted. This indicates that even in the absence of a visual field the upright body offers a stable basis for judging the V and H.

Another indication of the instability of perception when the body is tilted is the finding that in this condition judgments are even affected by the initial position of the rod (whether it is displaced in a clockwise or a conterclockwise direction). This is seen most strikingly in the 28° position.[4] When both body and rod were tilted 28° left, the mean error was 6.0°. When the body was 28° left, and the rod 28° right, the mean error rose to 9.3°. In both instances the most common reaction of the *S* was to bring the rod to a position on the other side of the upright from the body. It was displaced farther, however, when the rod was initially on the opposite side from the body than when it was on the same side. Not only the magnitude of the errors, but their direction as well, was determined by the initial position of the rod. When the rod was initially tilted with the body, it was displaced to the opposite side in 74.1 percent of all adjustments. Furthermore, 31 of the 49 *S*s tested displaced the rod to the opposite side in at least five of the six judgments made. On the other hand, when the rod was initially opposite to the body, it was 'sent the other way' in 81.0 percent of the judgments, and 36 *S*s adjusted the rod in this manner in at least five of their judgments. The fact that the magnitude and direction of the errors are influenced by the initial position of the rod indicates that these judgments do not have a very stable basis.

The above results show that when the body is tilted there is not a single position, but a range of positions, in which the rod appears upright. By starting the rod from either side of this range and then moving it in, as was done here, one may determine the peripheries of this range.

While the errors produced by tilting the head or body are very variable, in direction at least they are not random in character, but are systematically related to the position of the body. With *large* body tilts, the *S* perceived the V (or H) rod as displaced from the upright in the *opposite* direction from his body's actual displacement, and therefore moved it toward his body to make it appear V (or H). This is the Aubert or A-phenomenon (3). Conversely, with *small* body tilts, the upright rod was perceived as displaced *toward* the body, and was therefore moved away from it. This is the E-phenomenon, as named by Müller (7).[5] The direction of displacement of the rod at each body position is given in Table III. The A-phenomenon appeared most strikingly with the body in a horizontal position

[4] A similar effect was found for the 42° and horizontal positions of the body, though less strongly manifested.

[5] For a further account of these phenomena the reader is referred to the study by Müller (7).

(Experiment III). Here, in 360 out of 384 judgments (93.8 percent), the error was in the direction of the body tilt. The E-phenomenon was strongest at a 28° body tilt (Experiment IIb). Here, in 456 of 588 judgments (77.6 percent), the error was in the opposite direction from the body tilt. Of the remaining judgments, 8.0 percent were correct, hitting the V and H on the head, and 14.5 percent involved displacement of the rod toward the body. The E- and A-effects are manifested not only in the massed data but in individual performances as well. All but one of 30 *Ss* tested in the horizontal position displaced the rod toward the body in 75 percent or more of their judgments. With body tilted 28°, 38 of 49 *Ss* displaced the rod away from the body in 75 percent of their judgments. Only four displaced it toward the body. The remaining seven were more variable, failing to meet the 75 percent criterion.

TABLE III

DIRECTIONS OF ERRORS AT EACH TILT OF HEAD OR BODY

Body Position	Adjustment		
	Opposite Body	Toward Body	Brought to True V or H
Head alone tilted (Expt. I)	63.0%	30.4%	6.5%
Body tilted 28° (Expt. IIb)	77.6%	14.5%	8.0%
Body tilted 42° (Expt. IIa)	60.0%	31.0%	9.0%
Body horizontal (Expt. III)	5.2%	93.8%	1.0%

At body tilts between 28° and 90°, the direction of error in adjusting the rod is not so consistent as at these two extremes. With a 42° body tilt (Experiment IIa) the rod was displaced opposite to the body in 60.0 percent of all judgments, and toward the body in 31.0 percent of the judgments. In the remaining 9.0 percent, the V and H were correctly established. Considering individual performances, it is found that 13 of 30 *Ss* displaced the rod away from the body in 75 percent or more of their judgments, and four displaced it toward the body.[6] The remaining 13 *Ss* did not displace the rod consistently in either of these directions.[7] Both the massed results and the indi-

[6] The E-effect is stronger when the body is to the right than when it is to the left of the upright. With body tilted right, 75.0 percent of all adjustments were to the opposite side; with body left, only 45.0 percent. In terms of individual performances, 21 of 30 *Ss* displaced the rod to the opposite side in 75 percent or more of their judgments when tilted right; only eight of these 30 did so when tilted left. A similar left-right difference was found with the head alone tilted 45°. With head right, 69.6 percent of the judgments were to the opposite side; with head left, 56.5 percent. In terms of individual cases, 14 of 23 *Ss* consistently displaced the rod to the opposite side with head right, and only nine did so with head left. The basis of this left-right difference is not indicated by available evidence.

[7] Of these 13 *Ss*, five consistently displaced the rod to the left of the upright, regardless of the direction of body tilt. The basis of this kind of performance is not clear. Whatever factors

vidual results indicate that the 42° body position represents a region of transition from the E- to the A-effect, with neither as strong as at the 28° or the 90° position. While many *Ss* still show the E-effect, some few have already begun to show the A-effect. A great number, however, show neither. The existence of such differences suggests that individual factors may be important in determining the E- and A-effects, and in producing a shift from one to the other.

Just as the direction of errors depends on the position of the body, so does their magnitude. The results indicate that with body upright there are essentially no errors (average error, 1.5°). When the body is tilted from the upright, errors appear. Their magnitude is greater, however, at a 28° body tilt than at a 42° body tilt. For the former position the mean error was 7.6°, for the latter 5.9°. The errors were smaller for the 42° position because it is in the region of transition from the E- to the A-effect. In this region neither effect is strong, and many *Ss* are in the midst of shifting from a rod-opposite-body to a rod-toward-body type of adjustment. Beyond the 42° position, as the A-effect becomes stronger, the errors grow larger, reaching an average value of 16.1° for the horizontal position. A study which would employ additional body positions between the upright and the horizontal is required to establish more precisely the relation between body position and magnitude of errors.

With head or body tilted, there are significant individual differences in the magnitude of the errors made. At each position tested, with the exception of the body-erect position, *Ss* differed considerably in accuracy of approximation of the true V and H. The range of the results is best expressed by giving the poorest score and the best score for each test condition. These are as follows: head alone tilted, 1.9° and 12.3°; body tilted 28°, 1.9° and 20.8°; body tilted 42°, 1.3° and 14.3°; body horizontal, 6.9° and 37.6°. For each position, the range of individual scores is considerable, which suggests that in the absence of a visual field the tilted body is less useful for some people than it is for others in judging the upright.

A comparison of results for Experiment IIb, where the body was tilted 42°, and Experiment I, where the head alone was tilted about that amount, reveals a striking similarity in performance. In terms of magnitude of errors, the results are identical (5.9°) for both conditions. In terms of direction of errors, the results are again strikingly similar. The following gives the incidence of individual cases for each mode of rod adjustment, the first value referring to the body-tilted-42° condition and the second to the head-tilted condition: (a) errors to the opposite side of the V or H from the body, 43.3 percent

are responsible for this left displacement of the rod, they are completely submerged at the 28° and 90° body tilts. At these positions not a single *S* showed this type of adjustment.

and 47.8 percent; (b) errors in the direction of the body, 13.3 percent and 13.0 percent; (c) errors to the left of the V or H, 16.6 percent and 13.0 percent; (d) direction of errors variable, 26.6 percent and 21.7 percent.[8] The frequency of the different types of rod adjustment was thus about the same for the two conditions. Furthermore, the type of adjustment in which the rod was displaced to the left was observed only in these two body positions. Not a single instance of it occurred at the 28° or 90° position.

This striking similarity of performance when the head alone is tilted at a given angle and when the whole body is tilted at that angle suggests that the perceptual effects resulting from displacement of the body from the upright may have their main basis in the head region.

III. Discussion

When the visual framework is removed, it is no longer possible to judge the direction of a line by comparing it with other directional lines in the environment. Under such conditions, judgment must be based on postural factors. The experiments described here have demonstrated that, provided the body is *erect*, the determination of the V and H is very accurate. In this position the body is most useful, for all that is necessary is that the line being judged be brought into direct alignment with the felt position of the body. These findings are in keeping with those previously reported by other investigators. Neal (**8**) has demonstrated that *S*s when standing erect in a dark room, were very accurate in bringing a luminous rod to the upright. Jastrow (**4**) obtained similar results using a field which lacked directional V and H lines.

Whereas perception of the upright is not disturbed by removal of the visual framework when the body is erect, it is disturbed when the body is tilted or even when the head alone is tilted. This is probably because when the body is displaced, the task is no longer the simple one of aligning the rod with the felt position of the body; now some allowance must be made for its tilt. The disturbing effect of tilting the body in the absence of a visual field is indicated by three significant aspects of performance:

1. The H and V can no longer be established with accuracy. The errors vary in magnitude with the amount of body tilt, the largest errors occurring when the body is in a horizontal position. Individual factors also seem to affect the magnitude of the errors, so that in the same position some *S*s do much more poorly than others.

[8] In the head-tilted condition, there was a single case in which the adjustment was consistently to the right of the V or H. This type of adjustment did not occur at all in the body-tilted-42° condition.

2. Successive judgments made by the same individual are not consistent. Thus, in the course of a series of trials, the *S* perceived the luminous rod as upright not in a single position but over a range of positions. Such variability was not found in a previous study (2) where judgments were made in the presence of a visual framework. There, with the body tilted, judgments proved quite consistent, even though the available framework was a tilted one. In that study the *S* was tilted 22°, and his task was to adjust a rod to the V and H. The rod, however, was contained in a room. In one test, the room was tilted 22° with the *S*; in another, 22° to the opposite side from the *S*. Computing the range of each individual's scores and averaging the ranges, the following results were obtained for these two situations. When the room was tilted to the same side as the body, the mean range for the H was 2.0°, and the mean range for the V, 2.6°. When it was tilted to the opposite side from the body, the corresponding values were 3.6° [9] and 1.7°. These figures are quite small, indicating a high degree of consistency when judgments are made with reference to a visual field. In contrast, in the present dark-room experiments, when the *S* was tilted 28° but had no visual field available, the ranges were very large. The mean range for H judgments was 9.6°, the mean range for V judgments 10.8°. It may be concluded from these findings that the visual field provides a more stable frame of reference for judging direction than does the body.

3. A further consequence of tilting the body in the absence of a visual field is the increased difficulty encountered in making judgments. With an upright visual field present, judgments can be made almost instantaneously, and without any doubt as to their correctness. The same is most often true when the visual field is removed, provided the body remains erect. When the body is tilted in the absence of a field, however, some *S*s take a good deal of time to reach a decision, change their minds after arriving at a judgment, and when finally deciding on a position are unsure about it.

The errors produced by tilting the head or body are particularly striking in view of the nature of the changes involved. Although the body is tilted, the gravitational force acting upon it, which corresponds to the true upright, continues unchanged. Hence tilting the body does not remove the fundamental postural basis for judging the upright and at worst simply make it more difficult to detect the direction of the gravitational vertical. That so minor a change has such a disturbing effect points to the rather limited usefulness of postural determinants of the upright.

[9] The presence of one extreme case has exaggerated this figure. With that case eliminated, the mean range drops to 2.3°.

The same conclusion is suggested by the consequences of restoring the field while the body remains tilted. By simply turning on the light one immediately eliminates the errors described. All *S*s are then able to locate the V and H with complete accuracy, whether the body is tilted, horizontal, or even upside down. The visual field, when present, immediately assumes primacy over postural determinants. The latter become secondary, to the extent that changes in them which produced serious consequences in the absence of the visual field are now without effect.

From the results of the experiments described here it would of course not be correct to conclude that tilting the body in the absence of a visual field deprives a person of all basis for judgment. Even with body horizontal, most *S*s were able to judge the V and H within 20° of the true position of these directions. Twenty-two percent approximated the V and H within 10°, 43 percent within 15°, 70 percent within 20°, and 96 percent within 25°. Only one *S* was off by more than 25°, his mean error being 37.6°. The postural basis for judgment, however, although it permits a fair approximation of the upright when the body is tilted, is by no means an adequate one, and cannot be regarded as leading to proper orientation. In some situations—for example, in flying—errors of the order observed can lead to grave consequences.

Although the exclusion of a visual field is clearly responsible for the increased magnitude of errors and for their variability, it fails to explain the consistency in direction of the errors. If the errors had been due simply to the absence of a comparison 'standard' on which to base judgments, they would have varied randomly about the true V and H, instead of being consistently in one direction, as expressed by the E- and A-effects. Inasmuch as these effects occur only in the absence of a visual field, they must depend on special mechanisms which find expression upon removal of the field. In this regard they are related to a variety of other perceptual effects, such as the autokinetic phenomenon, and many illusions of size, depth, and distance, which are observed only when an 'anchoring' visual framework is lacking. In all these instances, specific and narrow factors (for example, such sensory features as the absolute size of the retinal image), which usually play a secondary role, emerge as important. It is not surprising therefore that phenomena such as the E- and A-effects should be unstable. Evidence of instability was found in the variations in successive judgments—not only in the magnitude of the errors but often in their direction as well. Because they occur in the absence of a framework, such phenomena are also influenced by subjective factors to a degree not possible under less ambiguous conditions. For example, Kleint (**5**) has found that through experimen-

tally induced attitudes it is possible to strengthen or eliminate the A-effect. Similarly, in the case of the autokinetic phenomenon, suggestion from another observer influences the direction and magnitude of the movement. Subjective factors prove effective in situations of this kind because here the observer must develop his own frame of reference, which is more readily influenced than a frame of reference derived from the objective situation.

The directional errors constituting the E- and A-effects seem to be determined by a variety of factors and to have a complex basis. They accordingly represent a large problem in their own right. Regardless of their origin, however, as noted by Koffka (**6**), they are significant because they provide a demonstration of the importance of the visual framework in perception. This demonstration comes mainly from the finding that the visual framework, when present, submerges and replaces as the main basis of perception a host of specific factors, which become important only when the framework is removed.

IV. Summary

In earlier studies of this series the effect of displacing the surrounding visual field upon perception of the upright was investigated. The present study was concerned with perception of the vertical and horizontal upon complete removal of the visual field. With the position of his body varied systematically, the *S*, in a completely darkened room, was required to adjust a luminous rod to the true vertical and horizontal. Judgments were found to be very accurate when the body was upright, but errors appeared as soon as the body, or even the head alone, was tilted. The largest errors occurred when the body was horizontal, which was the most extreme displacement of the body employed. In addition, when the body was tilted, successive judgments proved to be quite variable. It is thus indicated that postural factors provide an adequate and stable basis for judging the vertical and horizontal when the body is upright, but not when the body is tilted. The errors made with head or body tilted were systematic in their direction. With small tilts, the rod tended to be displaced opposite to the body, and with large tilts, toward the body. These are the long-known E-phenomenon and Aubert-phenomenon, respectively. These phenomena are eliminated at once upon introduction of a visual field.

(Manuscript received for immediate publication July 27, 1948)

References

1. Asch, S. E., & H. A. Witkin. Studies in space orientation: I. Perception of the upright with displaced visual fields. *J. exp. Psychol.*, 1948, **38**, 325–337.

2. Asch, S. E., & H. A. Witkin. Studies in space orientation: II. Perception of the upright with displaced visual fields and with body tilted. *J. exp. Psychol.*, 1948, **38**, 455–477.
3. Aubert, H. Eine scheinbare beteunde Drehung von Objekten bei Neigung des Kopfes. *Virchow's Arch.*, 1861, **20**, 381.
4. Jastrow, J. On the judgment of horizontal, vertical and oblique positions of lines. *Amer. J. Psychol.*, 1893, **5**, 220–223.
5. Kleint, H. Versuche über die Wahrnehmung. *Z. Psychol.*, 1937, **140**, 109–138.
6. Koffka, K. Perception, an introduction to Gestalt-Theorie. *Psychol. Bull.*, 1922, 19, 531–585.
7. Müller, G. E. Über das Aubertsche Phänomen. *Z. Psychol.*, 1916, II, **49**, 109–244.
8. Neal, E. Visual localization of the vertical. *Amer. J. Psychol.*, 1926, **37**, 287–291.

STUDIES OF TRACKING BEHAVIOR. I. RATE AND TIME CHARACTERISTICS OF SIMPLE CORRECTIVE MOVEMENTS

BY LLOYD V. SEARLE AND FRANKLIN V. TAYLOR

Naval Research Laboratory *

INTRODUCTION

As a result of increasing complexity in mechanisms of control, attention has been directed towards certain aspects of motor skill which are found to be of particular importance in the design of machines for human operation. A typical function is that of tracking, in which the operator is required to maintain alignment continuously between an indicator and some moving object or point, as in training a telescope, controlling an airplane, or driving an automobile.

The present experiments were undertaken to study the rate and time patterns which occur when an *S* is required to make corrective movements in response to suddenly presented error-displacements. For exploratory purposes, an experimental method is employed similar to that which the engineer uses to study the output error of mechanical transmission systems in response to step input functions.

Normally, the response of a mechanism under power control is not direct and instantaneous. It invariably involves certain time lags and is sometimes a fairly complex function of the control handle input movements. Similar statements can be made of the operator, who amplifies, modifies, and contributes time delays in the process of transferring data from his own input, represented by the visual presentation, to the control handle. An essential feature of the tracking situation is that the final output of the machine represents input to the tracker, while the tracker's output, in turn, becomes input to the machine. This circular situation is roughly continuous, and some degree of interaction necessarily results between response characteristics of the machine and those of the man.

To the engineer, it is important that the output of the machine is fed back into the machine as input, being modified in the process by the man's errors of tracking. When the man is replaced by an automatic tracking device (or some other electro-mechanical feedback) such a system consti-

* The writers are indebted to Dr. Alston S. Householder, Dr. Carl H. Wedell, Dr. Robert Y. Walker, and Mr. William W. Lambert for material contributions to apparatus and aid in formulating problems during early stages of these experiments.

The opinions or assertions contained herein are the private ones of the writers and are not to be construed as official or reflecting the views of the Navy Department or the naval service at large.

tutes a 'closed cycle control system,' or servomechanism. The operation of such a system is extremely complex, but methods have been developed for evaluating its performance in continuous operation when certain data are known regarding the transmission properties of its component mechanisms. These data, relating to power output, frequency relations, time lags, and amplitude distortion, are normally obtained by laboratory measurement, rather than by calculation, and by analyzing the outputs which occur in response to controlled step function or sine wave inputs. Functions of separate components are then combined in a series of differential equations. Certain assumptions must be made regarding additivity and linearity, but it has been found possible by using these equations to predict with reasonable accuracy the performance of the system in final operation.[1]

It is of interest that engineers have been led to this operational type of study for such systems, due to the very complexity in structure and function of the mechanical components. As Ellson (2) has pointed out, the method represents a direct parallel in engineering to the behavioral approach in psychology, wherein the basic measurements of input and output for a transmission system correspond with those of stimulus and response for an organism. A number of psychologists have suggested that the engineer's advance methods of quantitative analysis might be applied with profit to the study of human motor behavior. Such application would be especially useful in the case of skills in machine operation, since it would permit combining the operator's characteristics with those of the mechanical parts. The immediate difficulty with this proposal, however, arises from the basic assumptions of continuity, additivity, and linearity, which, although reasonable in the case of mechanical transmission systems, are known to be generally untrue of human response.[2] Craik (1), for example, has emphasized the notion that human operator response to momentary inputs is basically discontinuous, the apparent continuity of output under some conditions being a reflection of central control which integrates and extrapolates from past data. An increasing body of evidence tends to indicate that the engineer's present equations, complex as they are, are still far too simple to handle the human response relations adequately.

There is, nevertheless, a need for experimental data from which it may be determined, on the one hand, what specific extensions of present theory are required, and on the other, to what extent certain relations may be fitted into existing practice as useful approximations. The fact is well established that the operator's performance is better in some systems than in others. Accuracy and smoothness of operation are significantly affected, for example, when a change is made in gear ratio, inertia, or the aided tracking time constant. From this it may be inferred that the operator does have certain transmission characteristics which are relatively consistent and which, if

[1] A detailed explanation of essential principles of this technique, termed 'Operational Analysis,' is contained in the interesting paper by Ellson (2).

[2] The 'linearity' requirement in servo theory refers not only to linearity in the ordinary mathematical sense, but also implies (a) that response values are independent of time and (b) that the normal response at any time is a function only of past values of the input (3, p. 29). Thus, such factors as learning, fatigue, and anticipation constitute sources of nonlinearity in the human.

known, could be used to advantage in determining optimum characteristics for other parts of the system.

The experiments described below represent an initial, exploratory investigation of response to step input displacements of position. Interest is centered, in particular, on measures of reaction time, maximum rate, duration, and precision of corrective movements when these are treated as functions of the magnitude of presented displacements. Further considerations relate to the influence of individual differences, training, and modification of control with respect to sensitivity, inertia, and friction.

Apparatus

The *Ss'* corrective movements were recorded as pencil tracings of target lines which had been printed on flat sheets of paper. The apparatus employed was a device for moving the paper at constant speed behind a narrow horizontal slit, through which only a small segment of the target line was visible to the subject at any one instant. The sheets were 8½ in. by 14 in. in size, and

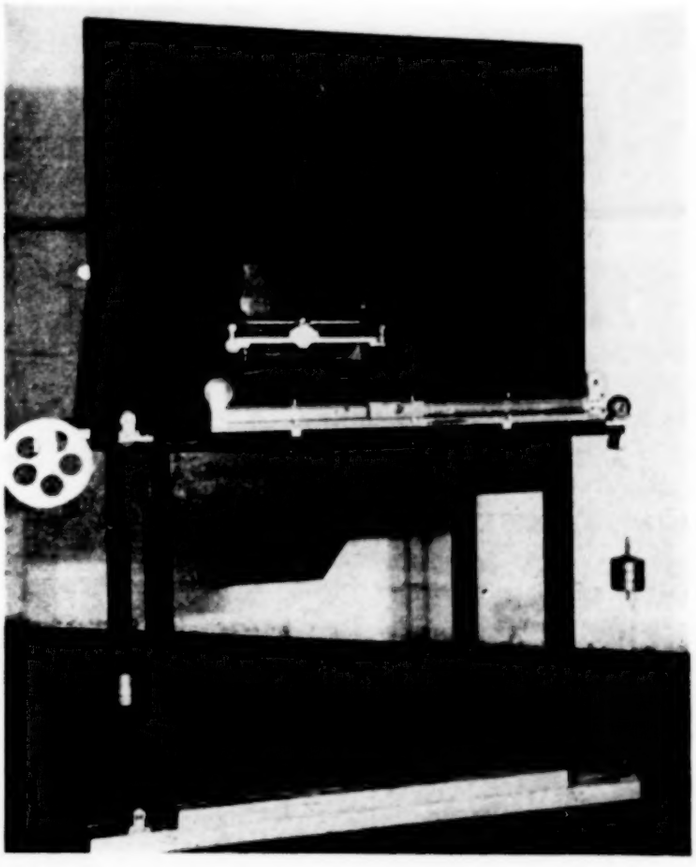

Fig. 1. Test apparatus as used in Experiments III and IV

were driven through the device lengthwise at the rate of 7½ cm. per sec. Thus, each sheet required approximately 4¾ sec. to run through. The target line, being printed as a straight line running lengthwise on the paper and having a jog in the middle, appeared through the slit to be stationary at one position on the sheet for about two or three sec., then to shift suddenly to one side and remain stationary in the new position.

The *S* 'tracked' by keeping the point of a pencil as nearly on the line as possible at all times. In the first experiments (Experiments I and II) this was done by holding the pencil in the hand. The *S* stood looking downward at the presentation and making his corrections by moving his hand in the horizontal plane either right-left or backward-forward, depending upon whether he stood in front of or to the side of the device.

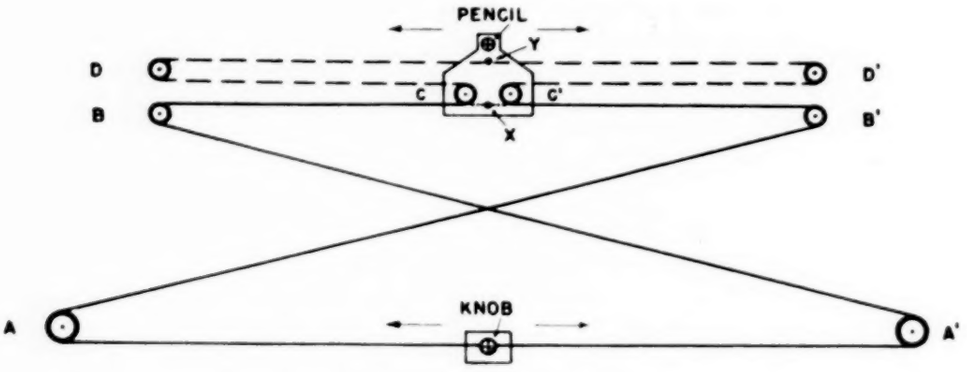

FIG. 2. Diagram of cord and pulley arrangement. Pulleys A, A', B, B', and D, D' are stationary; C and C' are mounted on the pencil slide bracket. For the 1:1 ratio of knob-to-pencil movement the cord is fastened at X, pulleys C, C' and D, D' not being used. To convert to 3:1 ratio the cord is replaced by a longer one which runs around the C, C' and D, D' pulleys and fastens at Y. In the latter case the pencil moves only one third as far as the knob in either direction.

A modification of the original device was made for subsequent experiments in order to standardize the pencil position relative to the paper for all *S*s and also to permit introducing changes in friction, inertia, and sensitivity of the control movement. The pencil was mounted on a horizontal slide and its movement was controlled remotely by means of a knob which was connected to the slide through a cord and pulley arrangement, as illustrated in Figs. 1 and 2. The arrangement permitted using either a 1:1 or a 3:1 ratio of knob-to-pencil motion. Also shown in Fig. 1 are the inertia weights and friction wheel, which were not connected to the knob slide except in the experiment dealing specifically with those variables (Experiment IV). The use of frictionless bearings and light materials throughout reduced the amounts of friction and inertia in the slide-pulley mechanism to negligible quantities. What little friction was present resulted from pressure of the pencil lead on the moving paper.

PROCEDURES

The various conditions studied have been summarized in the list of experiments which appears in Table I.

The purpose of the first experiment was to investigate the relation between size of the presented displacement and maximum rate of the corrective movement. Displacements presented to the *S*s required both right-going and left-going excursions of 5, 10, 20, 30, 40, 60, and 80 mm. in size. These were presented to each *S* in a series which included 10 repetitions of each size and direction (140 excursions in all), arranged in a predetermined random order. *S*s were instructed to hold the pencil point exactly on the printed line, making each correction as quickly and accurately as possible.

In the second experiment, *S*s stood to one side of the apparatus in such a position as to convert the left-right displacements into backward-forward (i.e., coming and going) movements of the hand in the horizontal plane. The pencil was held in the *S*'s hand in both of the first two experiments.

The apparatus was used in the form illustrated in Fig. 1 for Experiments III and IV. Instructions to the *S*s were similar, and random order series of presentations were used throughout. The number of *S*s, directions of presented displacements, and other conditions were varied as indicated in Table I.

TABLE I

CONDITIONS OF THE FOUR EXPERIMENTS

Experiment	Direction	Condition	Distance	No. of Each Size Displ.	No. of Subjects
I	Right	Subject holding pencil	5, 10, 20, 30, 40, 60, 80 mm.	10	4
	Left	Subject holding pencil	5, 10, 20, 30, 40, 60, 80 mm.	10	4
II	Forward	Subject holding pencil	5, 10, 20, 30, 40, 60, 80 mm.	10	4
	Backward	Subject holding pencil	5, 10, 20, 30, 40, 60, 80 mm.	10	4
III	Right	1:1 Ratio	5, 10, 20, 30, 40, 60, 80 mm.	20	5
	Right	3:1 Ratio	5, 10, 20, 30, 40, 60, 80 mm.	20	5
IV	Right	Control—(1:1 Ratio)	5, 10, 20, 30, 40, 60, 80 mm.	80	6
	Right	Friction—(1:1 Ratio)	5, 10, 20, 30, 40, 60, 80 mm.	40	6
	Right	Inertia —(1:1 Ratio)	5, 10, 20, 30, 40, 60, 80 mm.	40	6

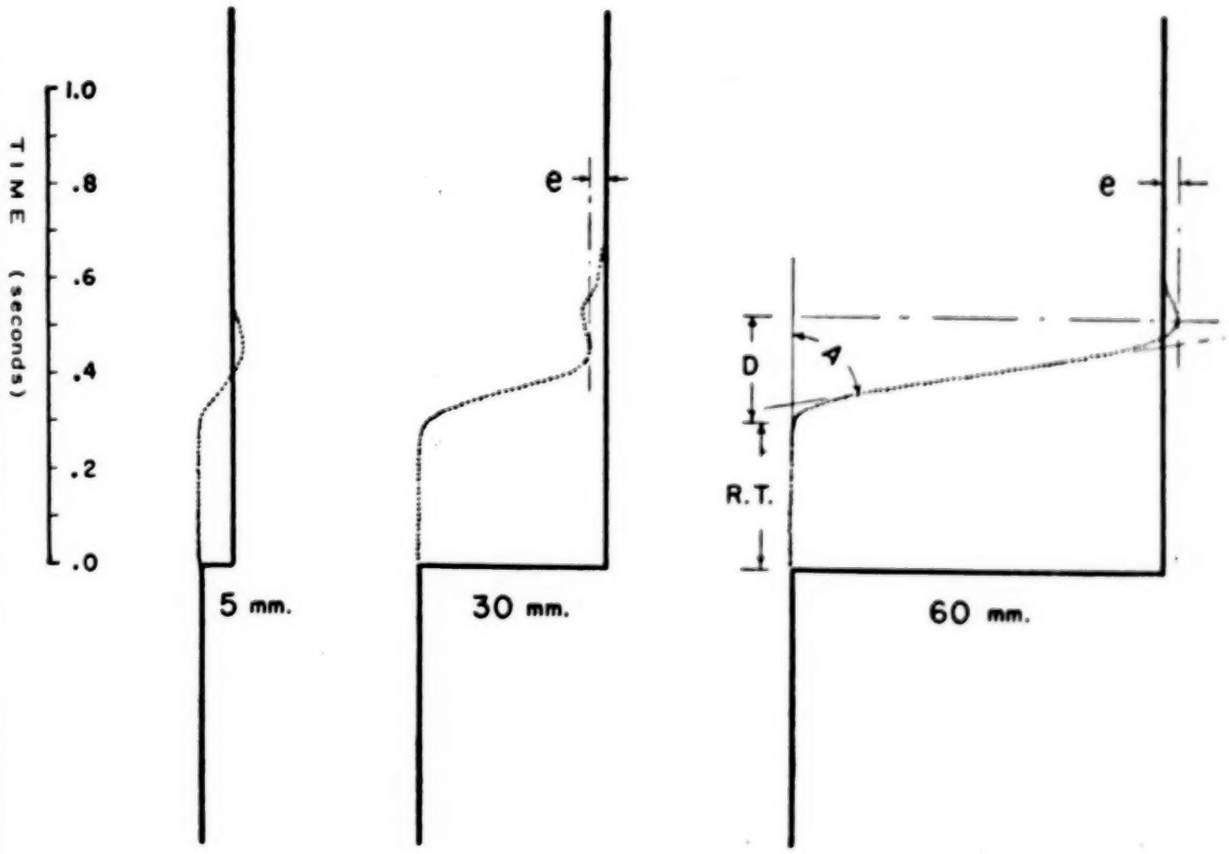

FIG. 3. *Reaction time* (R.T.) is measured as the distance between the occurrence of the presented displacement and the beginning of the corrective response. The *maximum rate* attained during the movement is computed as the slope (tangent) of angle A multiplied by the paper speed of 7.5 cm./sec. D is the *duration* of the response; e, measured in millimeters, is the *error*. The illustration shows 'overshoots' for the 5 and 60 mm., and an 'undershoot' in the case of the 30 mm. displacement.

Response Measures

Fig. 3 is a reproduction of the typical responses obtained to three sizes of displacements, the 5, 40, and 60 mm. sizes. The solid line in each case represents the printed target line, while the dotted lines are tracings of actual responses. A time scale has been added to the illustration representing lengthwise distances of the paper as it moved through the device at a constant rate of speed.

The *maximum rate* of speed reached during the course of the movement was obtained by the method which is indicated in the illustration of the 60 mm. excursion: the maximum slope of the line was determined by measuring its angular departure from the base, and this slope was multiplied by the longitudinal paper speed to obtain maximum rate in cm. per sec.

The *S*'s *reaction time* (RT) and the time from beginning to the completion of the response (D) were measured as distances on the paper and converted into seconds. The latter value, *duration*, represents an inverse measure of the average rate of movement.

Consideration was given in all cases only to the first, major corrective movement in reaching the new position. Emphasis was placed in the instructions to *S*s upon making the correction in a single movement, ending the movement as closely as possible to the new target line position. Hence, accuracy was measured as the difference (in mm.) between the target line and the point at which this first response ended. A statistical measure of *precision* was then computed as the standard deviation (i.e., variability) of these errors. Individual records in which the point of termination of the first response was unclear were omitted from the data analysis.

Results

Maximum rate vs. size of displacement

When the maximum rate measures are compared for the varying sizes of presented displacements, a clearcut tendency is found in the case of each *S* for correction rate to increase steadily as the magnitude of the 'error' increases. In Fig. 4 are shown the average curves for Experiments I and II, in which it may be seen that a definite positive relation exists between rate and size of displacement for all four directions of movement which were studied. It is evident by comparing

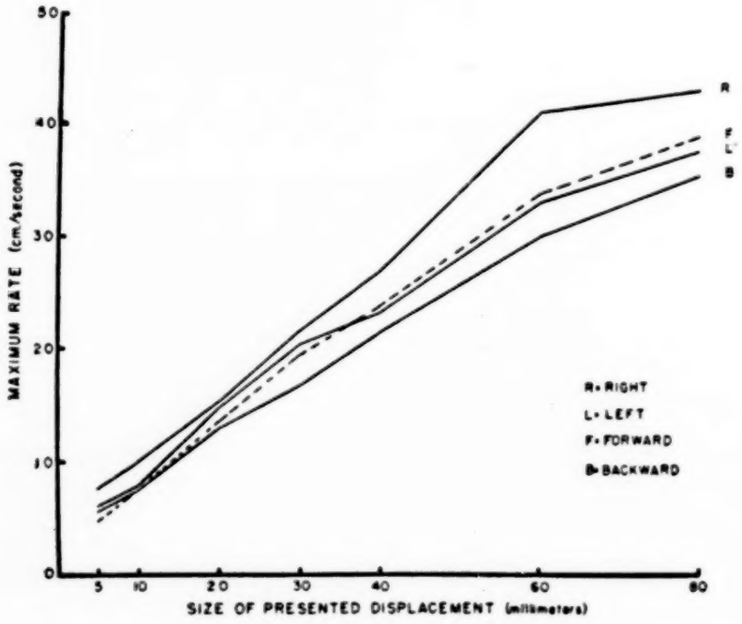

Fig. 4. Rate curves for four directions of movement

the four curves that rates are generally highest for right-going and forward movements, and lowest for movements in the backward direction. The increase in each case is approximately linear for distances up to 60 mm., but beyond that point the curves show a definite tendency to level off.

The latter finding suggest that a physiological limit is being approached in the correction rates which are reached at the 80 mm. movement. However, the variations among curves for individual *S*s, as well as variations for a given *S* with learning and other conditions, prevent drawing definite conclusions regarding the exact form of these rate curves. Each *S* differs from the others not only with respect to the absolute level of rates, but also with respect to the

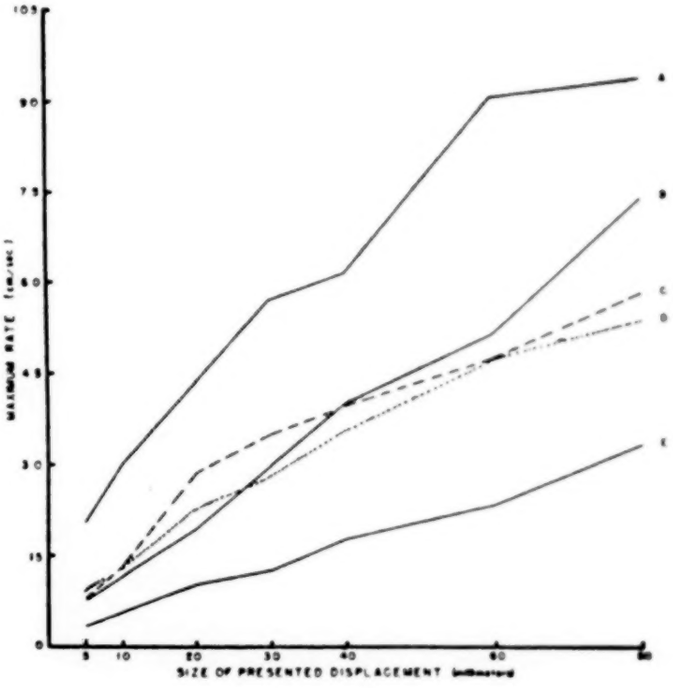

FIG. 5. Individual maximum rate curves for five *S*s

linearity of his curve. This is illustrated by Fig. 5, which shows the curves for the five *S*s of Experiment III. Of the five curves, two (B and E) are distinctly rectilinear, while the other three are negatively accelerated. A series of eight runs for a single *S* yielded the rate curves which are shown in Fig. 6. Here it is apparent that a given *S* shows some degree of inconsistency in his rates from trial to trial, both in form of the curve and average level. The variations appear to be related to a slight extent to the order in which the trials occurred, and therefore indicate some degree of learning. However, the fact that trials 7 and 8 show as much curvature, and as low rates at the 80 mm. excursion, as do trials 1 to 4 indicates that other sources of variation have a significant influence on the rates which are

made at each trial. These curves suggest that the *S*'s temporary 'set' for speed and accuracy are important in determining the form of his rate curve at each trial.

Two conclusions appear to be justified by the above results: (1) The maximum rate of a corrective movement shows a distinct and

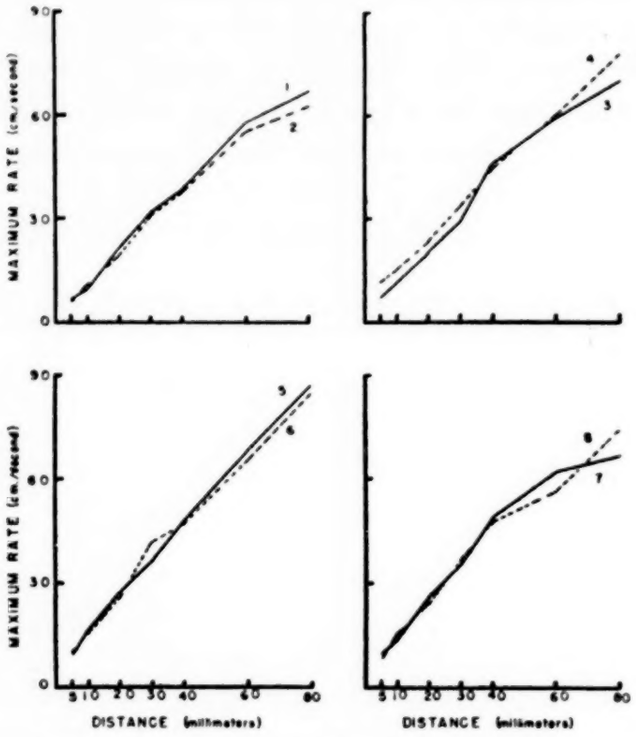

FIG. 6. Curves of maximum rate for one *S* on eight successive trials

systematic increase with increasing distances of the movement; (2) no exact form of this relation can be specified. It approximates a straight line in some cases, having a positive slope and positive-y-intercept, but often varies in the direction of becoming a negatively accelerated curve in the region of higher rates and longer excursions.

Reaction time

The average reaction time for all sizes of movements, computed for a total of 3360 responses by the six *S*s of Experiment IV, was found to be 0.257 sec. Reaction time shows no important variation as a function of displacement size, as may be observed in the upper solid line curve of Fig. 7, although the responses for 20 and 30 mm. appear to be slightly faster than others. A comparison of mean times yields the following results:

Mean of 5, 10, 40, 60, and 80:	0.262 sec.
Mean of 20 and 30:	0.245 sec.
Difference:	0.017 sec.
t-ratio:	6.82[3]

[3] Based on the six cases, this ratio represents a probability of less than .001 that the obtained difference would occur by chance.

The difference of 17 msec., therefore, indicates a slightly greater readiness on the part of all *Ss* to respond to these two displacement magnitudes. This results presumably, not from the absolute sizes of movements, but because 20 and 30 mm. represent middle values in the randomly presented series of sizes and, as such, were probably perceived on the average in less time. This 'range effect' has been noticed in connection with other measures and will be discussed further below. The indication is, in general, that *Ss* do not give comparable responses to the extreme high or low values in a presented series.

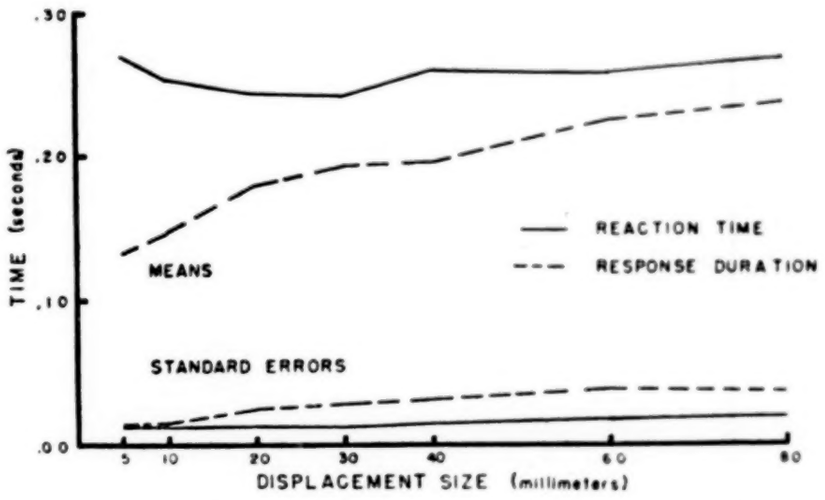

FIG. 7. Curves of response duration and reaction time, Experiment IV

Response duration

The mean values of response duration for the six *Ss* of Experiment IV are represented in the upper dashed line curve of Fig. 7. It may be seen that duration increases steadily as a function of the required hand movement. It is apparent, however, that the increase in time is by no means proportional to the increase in distance moved. Thus the 80 mm. movement, which is 16 times as great as the five mm., requires less than twice as long to make. This is consistent with the previously noted relation between maximum rate and distance, and implies that *average* rate would show a similar result. Average rate (computed from the measured response durations) and maximum rate are plotted together in the graph of Fig. 8 for purposes of comparison. Here, considering for the moment only the average rate curve, it may be noted that between the five mm. and 80 mm. excursions rate increases from 4 cm./sec. to 32 cm./sec., or about eight times. Thus, the average speed of movement does not increase in full proportion to the distance traversed, but it tends to do so much more nearly than does response duration. The evident tendency is for time of response to remain constant, regardless of distance, while

rate increases in proportion to the extent of movement. The slight but reliable increase in duration may be taken as indicating the degree to which complete constancy is not achieved.

Effect of Hand-to-Pointer Ratio

Results have been discussed in the preceding with reference only to the situation in which the 1:1 ratio of hand-knob to observed line and pencil movement was employed. In Experiment III a set of responses for each of the five Ss was also obtained at the 3:1 ratio, i.e., with the cord and pulley arrangement altered as shown in the diagram of Fig. 2.

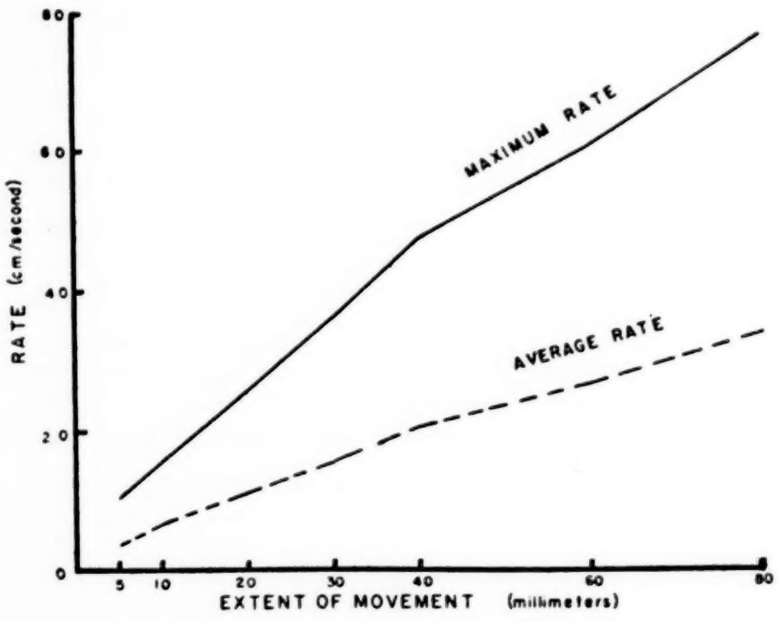

FIG. 8. Curves of average rate and maximum rate, Experiment IV

The experimental series was presented in the case of each *S* following a sample of practice excursions on the new ratio. No systematic study was made of learning during these practice trials, but it was noted that each *S* adapted very quickly to the changed condition, in most cases making an 'appropriate' response after only one or two trial movements. Presented displacements were identical to those employed with the 1:1 ratio. This meant that the observed pencil movements were equal in the case of both ratios (varying from five to 80 mm.), while the required hand movements were lengthened by exactly three times in the case of the 3:1 condition.

Results for the maximum rate measure are presented in the curves of Fig. 9. The solid line curve represents rate against displacement size for the 1:1 ratio and is identical with that shown in the preceding Fig. 8. The two dashed line curves drawn for the 3:1 ratio represent

rate against distance, in the one case, for rate and extent of movement at the hand, and in the other, for movement of the pointer. (The upper curve, therefore, is identical with the lower except that both ordinate and abscissal values are multiplied by three.)

Considering, first, the two lower curves, it may be seen that rates for the two ratios are essentially equal at the smaller excursions (below 30 mm.), while for the longer movements, 3:1 rates become progressively slower than the corresponding 1:1 rates. Since both of these

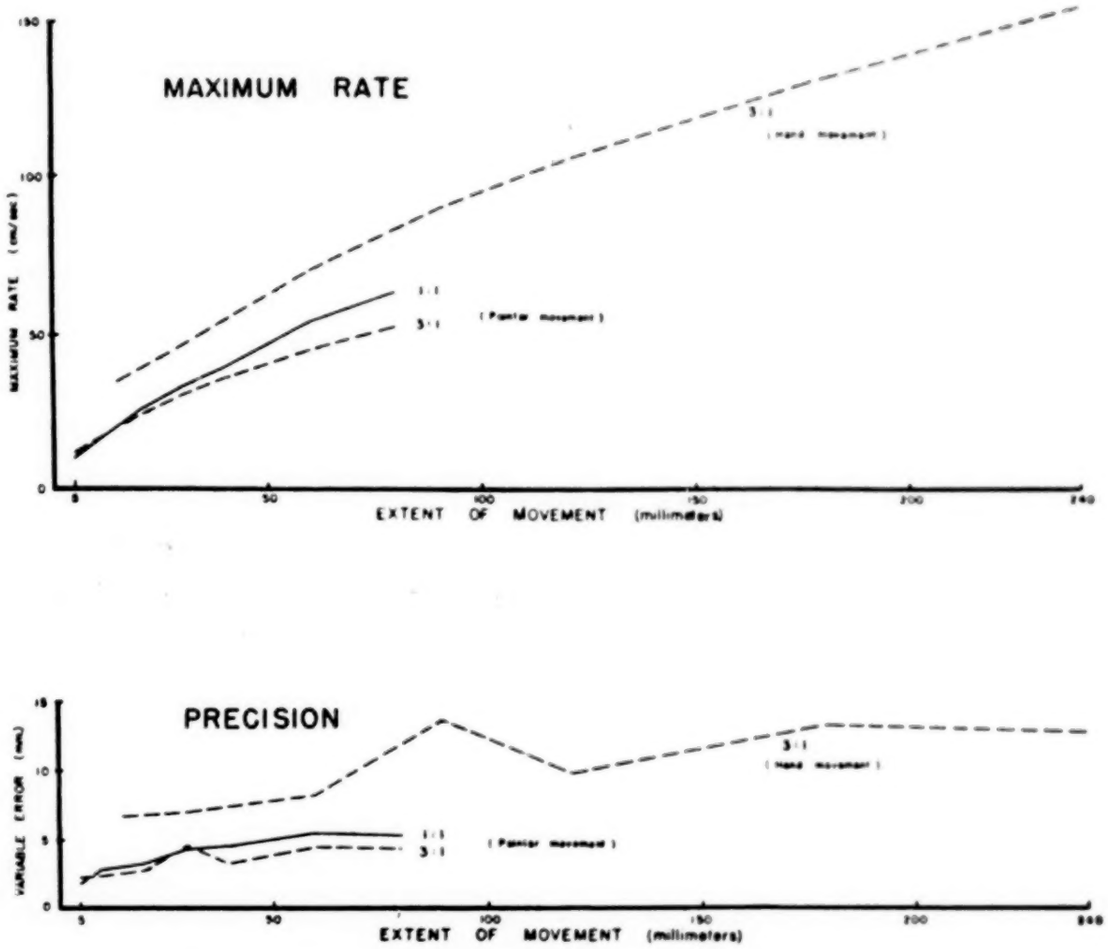

FIG. 9. Curves of maximum rate and precision, Experiment III

curves represent measurements taken directly from the printed sheets, their similarity indicates the extent to which the perceptual achievements in the two cases were alike. This is not altogether surprising, in view of the general finding that rate increases with distance, i.e., since an exact proportionality of rate with distance would cause all such curves to be superimposed, regardless of the hand-to-pointer ratio. A direct comparison of hand rates, however, indicates that the motor performance is not the same, as may be seen by comparing the 1:1 curve with the upper dashed line curve, which represents the

rates and distances of actual hand movement at the 3:1 ratio. Here, within the overlapping range of distances for which data were obtained, it is evident that the 3:1 rates were consistently higher.

A similar result occurred in the case of the measures of precision. As shown in the curves of Fig. 9, precision of observed pointer movement is consistently better at the 3:1 ratio than at the 1:1. The difference is relatively slight, however, when compared with the difference in precision of actual hand movement, in which the 3:1 ratio is definitely worse.

The present result clearly indicates that the actual level at which responses are made is influenced significantly by the perceptual aspects of the experimental situation, or by factors of internal set and expectancy which may vary with the conditions of presentation. Most directly, the increase in rates at the 3:1 ratio may be explained as a result of the reduced speed of perceived pencil movement, for which the *S*s involuntarily compensate by making faster hand movements. This explanation implies that the *S*s are attempting to make each correction in a certain length of time, regardless of the required effort, and that their increased speed is therefore a necessary means of preventing an undue increase in time of response. As was noted previously, response duration does change only slightly with wide variations in size of displacement.

There are other possible influences which should be considered. One of these is the 'range effect,' which was mentioned previously in connection with the reaction time measures. When displacements of varying sizes are presented in random order, *S*s' responses to any given size of displacement appear to be influenced by their relation to other sizes in the series, so that optimal responses occur to displacements which are somewhere near the middle in magnitude. Although the nature of this effect has not been investigated with any thoroughness in the present series of experiments, there are indications that rates are definitely too low at the upper end of a series, and too high at the lower end. The probable explanation appears to be that the *S*'s 'set' for responding to the average-size displacement influences all of his responses in the direction of the expected size, producing a clockwise rotation of the rate curve. It is difficult to estimate the amount of such an influence, but it seems likely that the higher level of 3:1 hand rates (in the region of movements from 15 to 80 mm.) may be accounted for at least in part by the greater range and average size of movements in the 3:1 series.

It was evident at times that the rates achieved were also influenced by the extent to which individual *S*s were differentially concerned with accuracy, although identical instructions were given to all *S*s throughout the series of experiments. The highest rates were made

by Ss who paid the least attention to terminating their excursions exactly on the line, and in the case of one or two Ss whose rates showed a temporary marked increase, the increase was accompanied by loss of precision. Other indications that accuracy is in general negatively correlated with rate are found in (a) the observation that accuracy decreases with increasing size of displacement (whereas rate increases), (b) the increase in accuracy with added inertia (where rate decreases), and (c) the present case of change in ratio, in which, as may be observed by comparing the two graphs in Fig. 9, the amounts of error and the rates achieved show the same relative patterns for the three conditions. In general, it is undoubtedly true that the S's own 'tolerance' for error exerts a limiting influence on his rate of movement, such that when he tries very hard to be precise his rate of movement naturally becomes slower. Since the curves of Fig. 9 indicate that Ss permit themselves approximately the same amount of *visual* error at both ratios, and since error increases only slightly with size of displacement, a given distance of hand movement is made with lower precision at the 3:1 ratio and can, consequently, be made at a higher rate.

Effects of friction and inertia

Data were obtained in Experiment IV under each of the following three conditions: (1) Control condition, in which minimum values of friction and inertia were present in the apparatus. These were determined by measurement to be 2.2 ounces and 5.8 ounces, respectively. (2) Friction condition, in which 18 ounces of friction was added by means of the wheel illustrated in Fig. 1. Attached to the shaft of the wheel at the rear is a dry disc type of clutch which can be adjusted to provide friction in varying amounts. The value specified, 18 ounces, was determined as the amount of running friction at the average rate of movement. Starting friction is slightly higher with this type of device, although the running friction was found to maintain a fairly constant value over widely varying rates of movement. (3) Inertia condition, in which 112 ounces of inertia was added by means of the weight and pulley arrangement shown in Fig. 1.

The amounts of additional friction and inertia in conditions (1) and (2) were chosen arbitrarily, since the purpose of these tests was largely to explore the nature of possible effects, rather than to determine the exact parameters. They were sufficient in each case to give a distinctly different 'feel' of the control knob as against the minimum conditions.

Data were taken only at the 1:1 ratio of knob-to-pointer movement. Two complete sets of responses were obtained from each of

the six *Ss* for inertia and friction conditions and four sets for the control condition. A set consisted of 20 responses to each of the seven sizes of presented displacements, all requiring right-going movements to correct.

Results are shown in Fig. 10, which presents the curves of response duration, maximum rate, and precision for each of the three

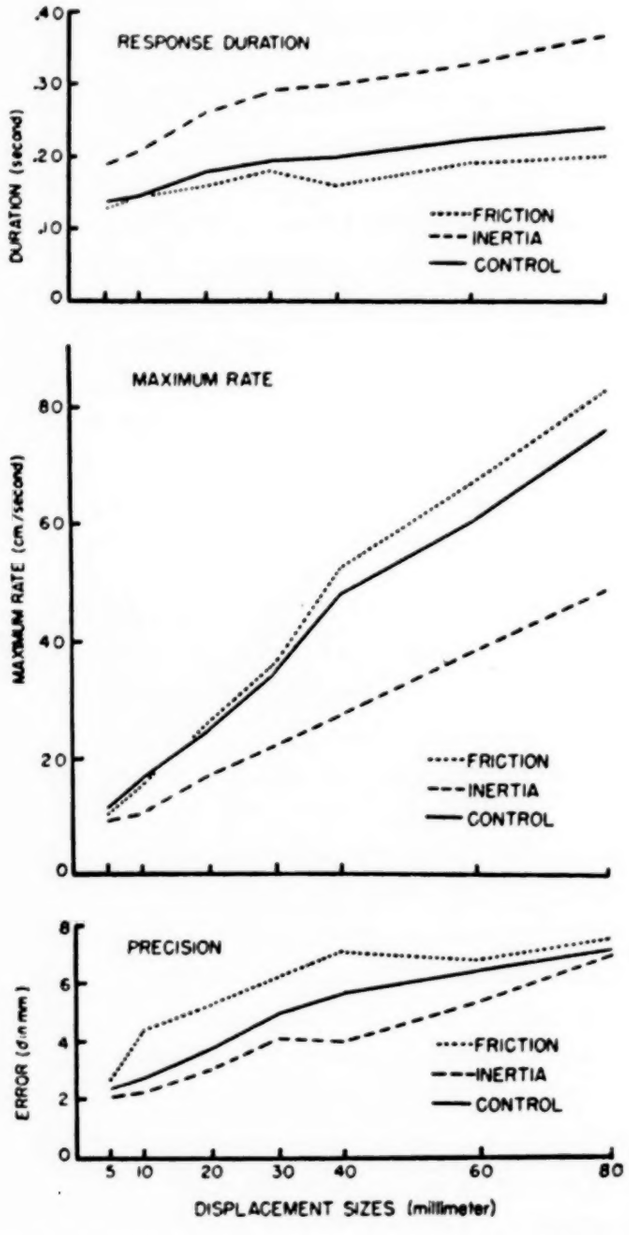

FIG. 10. Curves of response duration, maximum rate, and precision, Experiment IV

conditions. Referring to the curves, it may be observed that in general, friction and inertia have exactly opposite effects with respect to all three of the measures in question. Whereas inertia produces a decrease in rates and a slightly better precision, friction produces an increase in the rates and a decrease in precision of the movements.

These curves, although plotted from average values for the group of *Ss*, are generally typical of the effect upon individuals with only one or two minor exceptions. One *S*, for example, showed somewhat better precision with friction at the longer movements (30, 40, 60, and 80 mm.) than in the control series. Another *S* differed from the group as a whole in that his maximum rates were consistently lower with friction than without. The same *S*'s response durations, however, were shorter for the friction condition, indicating that his average rates and maximum rates were differentially affected. Since the response duration measure showed a more marked effect for other *Ss* as well, the inference may be drawn that friction acts in such a way as to change the acceleration characteristics of individual responses. Briefly, it may be assumed that the presence of friction reduces the overall response time by speeding up both the accelerating and decelerating phases of the movement, without necessarily increasing the maximum rate.

Discussion

The most consistent relation appearing throughout this series of studies is that between the rate and extent of a corrective movement. It may be stated as a general conclusion that rate increases with distance, since the relation has been found to hold for movements to the right, left, forward, and backward, for changes in sensitivity, and when appreciable amounts of friction or inertia are added to the control knob. No attempt has been made in the present study to investigate the causes of this relation, although it is of interest to inquire further into the nature of the sensori-motor principles involved. Questions arise as to whether, when errors of increasing extent are perceived, the resulting increase in rates of arm motion implies that differences must occur in the central organization of the response or whether, on the other hand, the increasing rates can be accounted for purely in terms of the physics of muscular movements. As regards the latter possibility, it is evident that to some degree, at least, a higher rate of arm motion can be developed if the accelerating force is applied for a longer period of time. It is known from the present experiment that response duration does increase, but whether or not the increase is sufficient to account for the marked change in rates cannot be determined in the absence of direct measures of the accelerations. A further study of this problem has already been made and will be reported in a succeeding paper (**4**).

However, it may be inferred from the observation that response durations are generally shorter than reaction times (Fig. 7) that a kind of intermittency must exist in the stimulus-response loop, since

each movement is brought to a complete stop before the visual signal of diminished error has had time to influence the stopping. Thus, as regards the human tracker's function as a servomechanism, there is indication that certain 'open-loop' phases occur during successive corrections in the tracking process, although it is apparent that an overall closed-loop situation exists. The indicated hypothesis is that during the interval of reaction time, the error is perceived, the response is organized, and an integrated temporal pattern of nerve impulses is triggered off. Judging from the present data, it seems unlikely that this pattern can be altered during its course in response to stimuli which occur after the pattern has been initiated.

Summary

Using a simple test apparatus, *Ss* were required to make arm-hand corrective movements of five to 80 millimeters in extent. Measures of reaction time, response duration, maximum rate, and accuracy are analyzed as a function of direction and extent of movement, sensitivity of the control knob, friction, and inertia.

1. The maximum rate and average rate of a movement increase systematically with increasing extent of the movement. The relation is approximately linear in most cases, although the increase is not strictly proportional because of a slight but reliable increase in response duration with the longer movements. Reaction times are relatively independent of the distance to be moved.

2. In the situations studied, rates are highest for movements from left to right and are successively lower for forward, left, and backward movements.

3. When control sensitivity is changed to a ratio of 3:1, the relations between rate and distance are altered with respect to both pointer and knob movements. For movements of comparable distance in each case, rates at the pointer are slightly lower, and rates at the hand are higher at the 3:1 than at the 1:1 ratio. In general, the operator tends to compensate for the change in sensitivity in such a way as to produce nearly the same perceived result with respect to both rate and accuracy.

4. The addition of an appreciable amount of inertia to the control knob causes the rates to be decreased for given distances of movement and the precision to be slightly improved. With additional friction, on the other hand, the responses are speeded up and precision is definitely lower than when minimum friction is present in the apparatus.

(Manuscript received for immediate publication July 15, 1948)

References

1. Craik, K. J. W. Theory of the human operator in control systems. I. The operator as an engineering system. *Brit. J. Psychol.*, 1947, **38**, 56–61.
2. Ellson, D. G. The application of operational analysis to human motor behavior. *Univ. Indiana Av. Psychol. Proj. Report No. 10.* Dayton: Aero Med. Lab., U.S.A.F., A.M.C. 3 May, 1948.
3. James, H. M., Nichols, N. B., & Phillips, R. S. *Theory of servomechanisms.* New York: McGraw-Hill, 1947.
4. Taylor, F. V., & Birmingham, H. P. Studies of tracking behavior. II. The acceleration pattern of quick manual corrective responses. *J. exp. Psychol.*, 1948, **38**. (In press)